



# Impact de la variabilité des données météorologiques sur une maison basse consommation. Application des analyses de sensibilité pour les entrées temporelles.

Jeanne Goffart

## ► To cite this version:

Jeanne Goffart. Impact de la variabilité des données météorologiques sur une maison basse consommation. Application des analyses de sensibilité pour les entrées temporelles.. Autre. Université Grenoble Alpes, 2013. Français. <NNT : 2013GRENA029>. <tel-00982150>

**HAL Id: tel-00982150**

**<https://tel.archives-ouvertes.fr/tel-00982150>**

Submitted on 23 Apr 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

### DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Génie Civil et Sciences de l'Habitat**

Arrêté ministériel : Arrêté ministériel : 7 Août 2006

Présentée par

**JEANNE GOFFART**

Thèse dirigée par **Etienne WURTZ**  
et codirigée par **Thierry MARA**

préparée au sein **LOCIE UMR CNRS 5271**  
et de **SISEO**

# Impact de la variabilité des données météorologiques sur une maison basse consommation. Application des analyses de sensibilité pour des entrées temporelles

Thèse soutenue publiquement le **12/12/2013**,  
devant le jury composé de :

**Mme Monika WOLOSZYN**

Professeur, Université de Savoie, Présidente

**M. Bertrand IOOSS**

Chercheur senior, EDF, Rapporteur

**M. Jean-Jacques ROUX**

Professeur, INSA Lyon, Rapporteur

**Mme Elena PALOMO DEL BARRIO**

Professeur, Université Bordeaux 1, Examinatrice

**M. Etienne WURTZ**

Directeur de Recherche, CEA-INES, Directeur de thèse

**M. Thierry MARA**

Maître de conférences, Université de la Réunion, Co-Directeur de thèse





# Remerciements

Il n'est pas évident de trouver les mots afin de retranscrire l'importance que certaines personnes ont pu avoir dans ma vie personnelle ou professionnelle durant ces trois années riches en rencontres et apprentissages.

Je souhaite tout d'abord remercier Etienne Wurtz, mon directeur de thèse, pour m'avoir offert l'opportunité de travailler sur une thématique aussi intéressante que la fiabilité et l'incertitude en thermique du bâtiment. J'ai fortement apprécié la confiance qu'il m'a témoigné à maintes reprises.

Je remercie Thierry Mara, mon co-directeur, de m'avoir accompagné dans l'apprentissage de l'analyse de sensibilité. Merci pour son accueil à la Réunion et pour les diverses discussions passionnantes que nous avons partagées.

Merci à mon jury de thèse de m'avoir fait l'honneur d'évaluer mes travaux. Merci à Jean-Jacques Roux et Bertrand Iooss d'avoir rapporté ma thèse. Il est important pour moi d'avoir pu bénéficier de leur expertise dans les domaines respectivement de la thermique du bâtiment et de l'analyse de sensibilité. Merci à eux, à Elena Palomo et Monika Woloszyn, ma présidente de jury, pour les discussions pertinentes durant la soutenance et pour l'intérêt qu'ils ont porté à mon travail.

Benoit Stutz, mon directeur de laboratoire, pour tous ses efforts pour rendre le LOCIE aussi convivial et pour son souci des autres. Merci également à Elisabeth, à Pascal pour sa présence et son dynamisme, à Jean-Patrick pour sa bienveillance. Merci à Monika qui a été un soutien et un moteur important durant ma dernière année de thèse. Mathieu, Yannick, Thierry, Parfait, Nina et autres membres du LOCIE pour la bonne ambiance que vous apportez.

L'équipe du LEB pour leur accueil et leur aide durant cette thèse, notamment à Jocelyn pour son ingéniosité, Frank pour son expertise en questionnaire de confort. Merci pour votre bonne humeur : Philippe, Arnaud, Patrice et le reste de l'équipe. Merci à Adrien, Louis, Layal, Clara : les "anciens", pour leurs conseils. Merci à mes co-bureaux pour les moments partagés : Xavier, Jonathan, Yannick et les autres. Merci à Isabelle pour l'intérêt qu'elle porte aux personnes qu'elle côtoie.

La thèse est une aventure humaine et j'ai eu la chance d'y avoir partagé mon quotidien avec des gens extraordinaires que je souhaite remercier pour leur présence sans laquelle cette aventure aurait été moins enrichissante. Merci à Fabio et Gianpiero de m'avoir fait découvrir leur Sicile et d'avoir pris autant soin de moi. Merci à Malik et Vincent parce qu'être à la Réunion "c'est pas facile" . A Aurélie compagne des différentes missions au bout du monde. A Lilian pour nos tests de "pointillé-craie". Yann et Madi "la prince" pour leur douceur à mon égard. Julien "le breton" pour nos bières au bord du lac. A Seb, Mel et Mickey pour avoir été mes compagnons de vie durant cette aventure et pour avoir réussi à recréer un "chez moi".

Un grand merci à Coralie et Sophie pour être là et pour leur investissement dans un des meilleurs pots qu'il m'a été donné d'assister (et évidemment je reste objective).

Enfin, à ma famille, qui a toujours cru en moi, quelque soit mon projet. Merci à mes parents pour leur soutien inconditionnel, à mes frères : Hugues et Mathieu. Obrigada a minha familia brasileira : Carlos, a Marcia e a Luciana. Merci aussi à Antoine et Rosario, mes grand-parents espagnols. **Votre fierté me fait avancer.**

Merci à tous ces gens qui m'ont témoigné leur soutien dans les moments difficiles mais également dans les réussites.

# Tables des matières

Remerciements	i
Liste des tableaux	iv
Table des figures	viii
Nomenclature	ix
<b>1 Contexte de l'étude et définition de la problématique : Incertitudes et bâtiment à haute efficacité énergétique</b>	<b>3</b>
1.1 De la gestion des incertitudes en simulation thermique dynamique vers la garantie de performance . . . . .	4
1.1.1 Nécessité d'un gain en fiabilité . . . . .	4
1.1.1.1 La modification des postes de dépense énergétique . . . . .	4
1.1.1.2 La simulation en réponse aux attentes des concepteurs . . . . .	5
1.1.2 Les incertitudes : concept et terminologie . . . . .	6
1.1.2.1 Comment définir une grandeur physique? . . . . .	6
1.1.2.2 Quelques notions sur l'incertitude . . . . .	6
Incertitude aléatoire et incertitude systématique . . . . .	6
Exactitude et précision . . . . .	6
Notion d'intervalle et de niveau de confiance . . . . .	8
1.1.3 Discussion sur la présence des incertitudes en simulation thermique dynamique . . . . .	10
1.1.3.1 Vers des logiciels de plus en plus complexes . . . . .	10
1.1.3.2 Les incertitudes aux différentes étapes de la simulation thermique dynamique . . . . .	12
Les données d'entrée . . . . .	12
Le modèle . . . . .	13
La sortie et son interprétation . . . . .	13
L'utilisateur . . . . .	13

1.2	Exemple de sources d'incertitudes et de leurs impacts sur une maison basse consommation . . . . .	15
1.2.1	Présentation du cas d'étude : Maison INCAS DM . . . . .	15
1.2.1.1	Description du bâtiment test . . . . .	15
1.2.1.2	Description du modèle associé . . . . .	17
	Le logiciel de simulation EnergyPlus . . . . .	17
	Hypothèses de simulation . . . . .	18
1.2.2	Impact de l'incertitude de l'albédo . . . . .	19
1.2.2.1	Description de l'entrée . . . . .	19
1.2.2.2	Un paramètre incertain et variable . . . . .	19
1.2.2.3	Résultats de l'impact sur la maison INCAS-DM . . . . .	21
1.2.3	Impact d'une hypothèse de simulation sur les masques proches . . . . .	22
1.2.3.1	Description des débords de fenêtre . . . . .	22
1.2.3.2	Impact de l'hypothèse de simulation sur la maison INCAS-DM . . . . .	22
1.3	Un besoin de transparence sur la gestion des sollicitations par le bâtiment basse consommation . . . . .	24
<b>2</b>	<b>Outils Mathématiques pour le gain en fiabilité : les analyses de sensibilité globale et d'incertitude</b>	<b>27</b>
2.1	Des outils puissants . . . . .	27
2.1.1	Rayons X du modèle conditionnés par un objectif . . . . .	27
2.1.2	Plage de variation et zone d'exploration du modèle : cas de la thermique du bâtiment . . . . .	29
2.1.3	Choix de la méthode . . . . .	31
2.2	Des indices de sensibilité adaptés à la problématiques du bâtiment . . . . .	33
2.2.1	Des indices basés sur la variance . . . . .	33
2.2.2	Démarche et bonnes pratiques pour l'obtention des indices . . . . .	34
2.2.2.1	Échantillonnage par hypercube latin . . . . .	35
2.2.2.2	Une estimation précise . . . . .	36
2.2.3	Outils complémentaires : intervalle de confiance et réduction du coût de simulation pour les $S_i$ . . . . .	38
2.2.3.1	Réduction drastique du nombre de simulation par permutation de matrice . . . . .	38
2.2.3.2	Quantification de la variabilité sur les indices de sensibilité par bootstrap . . . . .	39
2.3	Gestion des entrées fonctionnelles en analyse de sensibilité . . . . .	41
2.3.1	Une problématique actuelle en thermique du bâtiment . . . . .	41
2.3.2	Incertaines associées aux fichiers météorologiques . . . . .	42

2.3.3	Un besoin de méthodologie qui dépasse le domaine de la thermique du bâtiment . . . . .	44
<b>3</b>	<b>Méthodologie pour la prise en compte des données météorologiques dans les analyses de sensibilité</b>	<b>47</b>
3.1	Méthode de génération développée . . . . .	47
3.1.1	Principe de la méthode . . . . .	47
3.1.1.1	Génération du champ aléatoire $\varepsilon_i(\theta, t)$ par Iman et Conover	48
3.1.1.2	Génération du champ aléatoire $\varepsilon_i(\theta, t)$ par Karhunen-Loève	49
3.2	Application de la méthode sur des conditions climatiques . . . . .	50
3.2.1	Conditions initiales de la génération . . . . .	50
3.2.1.1	Le fichier source . . . . .	50
3.2.1.2	La période source . . . . .	51
3.2.2	Extraction des données statistiques . . . . .	53
3.2.3	Génération des échantillons . . . . .	57
3.3	Vérification de la cohérence des échantillons . . . . .	59
3.3.1	Dispersion et accord avec le fichier source . . . . .	59
3.3.2	Évolution de l'inter-corrélation durant le processus de génération . .	61
3.3.3	Couverture de l'espace des paramètres . . . . .	66
3.3.4	Comparaison aux normales mensuelles de Météo-France . . . . .	69
3.3.4.1	Extraction des normales mensuelles des échantillons générés	70
3.3.4.2	Résultats pour Lyon au mois de juillet . . . . .	71
3.3.4.3	Résultats pour Lyon au mois de janvier . . . . .	74
3.4	Application à une autre localisation : la ville de Strasbourg . . . . .	78
3.4.1	Normales mensuelles de Météo-France . . . . .	78
3.4.2	Comparaison des échantillons . . . . .	78
3.4.2.1	Résultats pour Strasbourg au mois de janvier . . . . .	78
3.4.2.2	Résultats pour Strasbourg au mois de juillet . . . . .	80
3.5	Des échantillons cohérents : vers l'analyse de sensibilité . . . . .	82
<b>4</b>	<b>Estimation de l'incertitude des besoins énergétiques d'un bâtiment liée à la variabilité météorologique</b>	<b>85</b>
4.1	Protocole . . . . .	85
4.1.1	Objectif et contexte de l'étude . . . . .	85
4.1.2	Description de la méthode . . . . .	86
4.1.3	Analyse des sorties de simulation . . . . .	88
4.1.3.1	Le besoin de chauffage et de refroidissement . . . . .	88
4.1.3.2	Analyse des pertes et des gains par vitrage . . . . .	88
4.2	Comparaison de Lyon et Strasbourg : impact du climat étudié . . . . .	89

4.2.1	Réponse du bâtiment en période hivernale . . . . .	89
4.2.1.1	Résultats obtenus pour le climat de Lyon . . . . .	89
	Dispersion des résultats . . . . .	89
	Indice de sensibilité sur les besoins de chauffage : les variables responsables de la dispersion . . . . .	94
	Indices de sensibilité sur les gains et pertes par vitrage : vé- rification de la cohérence de la sensibilité . . . . .	95
	Conclusion . . . . .	97
4.2.1.2	Comparaison des résultats obtenus avec le climat de Stras- bourg . . . . .	98
	Dispersion des résultats . . . . .	98
	Indice de sensibilité . . . . .	101
4.2.1.3	Conclusion . . . . .	104
4.2.2	Réponse du bâtiment en période estivale . . . . .	104
4.2.2.1	Résultats obtenus pour le climat de Lyon . . . . .	104
	Dispersion des résultats . . . . .	104
	Indices de sensibilité sur les dispersions . . . . .	107
4.2.2.2	Résultats obtenus pour le climat de Strasbourg . . . . .	109
	Dispersion des résultats . . . . .	109
	Indice de sensibilité . . . . .	110
4.2.3	Conclusion . . . . .	111
4.3	Comparaison de modèles pour le même climat . . . . .	112
4.3.1	Résultats pour le cas "Triple vitrage" . . . . .	112
4.3.1.1	Protocole . . . . .	112
4.3.1.2	Estimation en période hivernale . . . . .	112
4.3.1.3	Estimation en période estivale . . . . .	115
4.3.2	Résultats pour un bâtiment moins performant : cas "Maison Dégradée" . . . . .	117
4.3.2.1	Protocole . . . . .	117
4.3.2.2	Résultats de l'étude "Maison Dégradée" pour le mois de janvier . . . . .	118
4.3.2.3	Résultats de l'étude "Maison Dégradée" pour le mois de juillet . . . . .	120
4.3.3	Comparaison de la variabilité temporelle des indices : effet du sto- ckage sur les indices de sensibilité . . . . .	122
4.4	Variabilité et sensibilité du modèle pour des entrées statiques et temporelles . . . . .	127
4.4.1	Analyses de sensibilité et d'incertitude pour les entrées statiques . . . . .	128
4.4.1.1	Résultats de l'étude statique pour le mois de janvier . . . . .	128
4.4.1.2	Résultats de l'étude statique pour le mois de juillet . . . . .	129

---

4.4.2	Analyses de sensibilité et d'incertitude avec des entrées temporelles et statiques . . . . .	130
4.4.2.1	Résultats de l'étude mixte des entrées statiques et temporelles pour le mois de janvier . . . . .	130
4.4.2.2	Résultats de l'étude mixte des entrées statiques et temporelles pour le mois de juillet . . . . .	131
4.5	Conclusion et tableau des dispersions . . . . .	132
	<b>Conclusion</b>	<b>135</b>
	<b>Perspectives</b>	<b>137</b>
	<b>Annexes</b>	<b>145</b>
	Plan et composition de la maison INCAS . . . . .	146
	Données statistiques extraites du fichier IWECC Lyon au mois de juillet . . . . .	148
	Valeurs et dispersions des 49 paramètres statiques de l'analyse de sensibilité statique et mixte . . . . .	154
	<b>Résumé</b>	<b>157</b>
	<b>Abstract</b>	<b>159</b>



# Liste des tableaux

1.1	Probabilité Pr (en %) pour que la valeur d'une variable gaussienne $x$ soit dans l'intervalle $[\mu - r\sigma, \mu + r\sigma]$ pour diverses valeurs de $r$ . . . . .	9
1.2	Variabilité de l'albédo en fonction de la nature du sol . . . . .	19
1.3	Mesures préliminaires sur site de l'albédo . . . . .	20
3.1	Valeurs des coefficients de corrélation de Pearson des six variables pour le signal source . . . . .	64
3.2	Données des normales mensuelles Météo-France en janvier et juillet de 2004 à 2013 pour la station Lyon St Exupéry . . . . .	70
3.3	Moyenne et écart type des échantillons et intervalle de confiance au seuil de probabilité de 95% . . . . .	72
3.4	Écart type, moyenne des échantillons et intervalle de confiance au seuil de probabilité de 95% . . . . .	75
3.5	Données des normales mensuelles Météo-France de 2004 à 2013 pour la ville de Strasbourg . . . . .	78
3.6	<b>Strasbourg en janvier</b> Moyenne et écart type des échantillons et intervalle de confiance au seuil de probabilité de 95%. . . . .	79
3.7	<b>Strasbourg en juillet</b> Moyenne, écart type des échantillons et intervalle de confiance au seuil de probabilité de 95%. . . . .	81
4.1	<b>Lyon en janvier</b> Valeurs de dispersion du besoin de chauffage total et par zone . . . . .	92
4.2	<b>Strasbourg en janvier</b> Valeurs de dispersion du besoin de chauffage total et par zone . . . . .	99
4.3	<b>Lyon en juillet</b> Valeurs de dispersion du besoin de refroidissement total et par zone . . . . .	104
4.4	<b>Strasbourg en juillet</b> Valeurs de dispersion du besoin de refroidissement total et par zone . . . . .	109
4.5	<b>Triple vitrage en janvier</b> Valeurs de dispersion du besoin de chauffage total et par zone . . . . .	113

4.6	<b>Triple Vitrage en janvier</b> Valeurs des indices de sensibilité $S_i$ et de son estimé bootstrap $\hat{S}_i$ pour le besoin de chauffage total et de chaque zone . . .	115
4.7	<b>Triple vitrage en juillet</b> Valeurs de dispersion du besoin de chauffage total et par zone . . . . .	116
4.8	<b>Triple Vitrage en juillet</b> Valeurs des indices de sensibilité $S_i$ et de son estimé bootstrap $\hat{S}_i$ pour le besoin de chauffage total et de chaque zone . .	117
4.9	<b>Modèle Dégradé en janvier</b> Valeurs de dispersion du besoin de chauffage total et par zone . . . . .	118
4.10	<b>Maison Dégradée en janvier</b> Valeurs des $S_i$ et $\hat{S}_i$ pour le besoin de chauffage total et de chaque zone . . . . .	120
4.11	<b>Maison Dégradée en juillet</b> Valeur de dispersion du besoin de chauffage total et par zone . . . . .	121
4.12	<b>Maison Dégradée en juillet</b> Valeurs des indices de sensibilité $S_i$ et de son estimé bootstrap $\hat{S}_i$ pour le besoin de chauffage total et de chaque zone	122
4.13	<b>Statique en janvier</b> Valeurs de dispersion du besoin de chauffage total et par zone . . . . .	128
4.14	<b>Statique en janvier</b> Variables les plus influentes et valeurs des indices de sensibilité $S_i$ pour le besoin de chauffage total et de chaque zone . . . . .	129
4.15	<b>Statique en juillet</b> Valeur de dispersion du besoin de refroidissement total et par zone . . . . .	129
4.16	<b>Statique en juillet</b> Variables les plus influentes et valeurs des indices de sensibilité $S_i$ pour le besoin de refroidissement total et de chaque zone . . .	130
4.17	<b>Statique et dynamique en janvier</b> Valeurs de dispersion du besoin de chauffage total et par zone . . . . .	130
4.18	<b>Statique et Dynamique en janvier</b> Variables les plus influentes et valeurs des indices de sensibilité $S_i$ pour le besoin de chauffage total et de chaque zone . . . . .	130
4.19	<b>Statique et Dynamique en juillet</b> Valeurs de dispersion du besoin de refroidissement total et par zone . . . . .	131
4.20	<b>Statique et Dynamique en juillet</b> Variables les plus influentes et valeurs des indices de sensibilité $S_i$ pour le besoin de refroidissement total et de chaque zone . . . . .	131
4.21	Récapitulatif des dispersions pour les différents cas . . . . .	133

# Table des figures

1.1	Illustration des différentes configurations de précision et d'exactitude par analogie avec une cible et le type de gaussienne correspondante . . . . .	7
1.2	Représentation sur la distribution de gauss des probabilités de trouver $x$ dans un intervalle de $2\sigma$ (68,27%), $4\sigma$ (95,45%) et $6\sigma$ (99,73%) . . . . .	9
1.3	Représentation de l'incertitude potentielle dans les prédictions de performance en STD en fonction de la complexité du modèle [67] . . . . .	11
1.4	Décomposition schématique de la simulation thermique dynamique . . . . .	12
1.5	Principaux matériaux composant l'enveloppe du cas test INCAS-DM [61] .	16
1.6	Prise de vue de la maison DM associée aux croquis de la distribution des vitrages en fonction des 4 orientations. . . . .	17
1.7	Hierarchie et besoin d'EnergyPlus . . . . .	18
1.8	Diagramme de répartition des températures intérieures dans la maison INCAS-DM sur la période estivale en fonction des différents cas [18] . . . . .	21
1.9	Les captures écran des deux cas simulés pour les fenêtres. A gauche le cas A : casquette seule et à droite le cas B : débords complets . . . . .	23
1.10	Type de masque proche sur les fenêtres pour le cas A : casquette seule et le cas B : débords complets . . . . .	23
1.11	Diagramme de répartition des températures intérieures dans la maison INCAS-DM sur la période estivale en fonction des différents cas, en bleu foncé le cas A casquette seule, en bleu clair le cas B débords complets [18] . . . . .	24
2.1	Evolution de la variable de sortie inconfort d'hiver en fonction de la variabilité générée de l'entrée "épaisseur d'isolant" . . . . .	30
2.2	Illustration du principe de l'échantillonnage hypercube latin . . . . .	36
3.1	Évolution de la température du fichier moyen (IWEC) de Lyon . . . . .	52
3.2	Besoins horaires de la maisons INCAS-DM sur l'année pour Lyon . . . . .	53
3.3	Procédure d'extraction à partir du signal concaténé . . . . .	55
3.4	Procédure d'extraction à partir des valeurs horaires sans la moyenne . . . . .	56
3.5	Exemple de 20 échantillons obtenus parmi les 1000 pour les six variables météo . . . . .	58

3.6	Dispersion des valeurs moyennes journalières pour les 1000 mois de l'échantillon et les six variables météo pour Lyon au mois de juillet . . . . .	60
3.7	Corrélation entre la température et l'humidité extérieures aux différentes étapes de génération, pour un jeu d'échantillon tiré aléatoirement parmi les $N$ . . . . .	62
3.8	Comparaison d'un échantillon et du signal source pour la corrélation entre la température et l'humidité extérieures . . . . .	63
3.9	Représentation en matrice des inter-corrélations pour les six variables . . .	65
3.10	Comparaison du recouvrement de l'espace des paramètres par scatterplot de 10 échantillons parmi les 1000 et du signal source . . . . .	68
3.11	Dispersion des 1000 valeurs pour les quatre normales mensuelles sous forme de boxplot . . . . .	71
3.12	Écart à la loi normale pour les normales mensuelles des échantillons de Lyon en juillet . . . . .	72
3.13	Comparaison des observations de Meteo-France par rapport à la dispersion des échantillons : Lyon en juillet . . . . .	74
3.14	Dispersion des 1000 valeurs pour les quatre normales mensuelles par boxplot	75
3.15	Écart à la loi normale pour les normales mensuelles des échantillons . . . .	76
3.16	Comparaison des observations de Meteo-France par rapport à la dispersion des échantillons : Lyon en janvier . . . . .	77
3.17	Comparaison des observations de Meteo-France par rapport à la dispersion des échantillons : Strasbourg en janvier . . . . .	80
3.18	Comparaison des observations de Météo-France par rapport à la dispersion des échantillons : Strasbourg en juillet . . . . .	82
4.1	Récapitulatif des étapes pour l'évaluation de la sensibilité et de l'incertitude d'un modèle de bâtiment aux sollicitations météorologiques . . . . .	87
4.2	<b>Lyon en janvier.</b> Dispersion des 2000 valeurs de besoins de chauffage . .	90
4.3	<b>Lyon en janvier.</b> Écart à la normale pour les besoins de chauffage . . . .	91
4.4	<b>Lyon en janvier.</b> Dispersion sous forme de boxplot des pertes et gains par vitrage pour chaque zone et orientation. . . . .	93
4.5	Distribution des surfaces vitrées de la maison INCAS-DM en fonction des orientations . . . . .	94
4.6	<b>Lyon en janvier</b> Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total . . . . .	95
4.7	<b>Lyon en janvier.</b> Indice de sensibilité principal sur chaque orientation pour les pertes par les fenêtres de chaque zone . . . . .	97
4.8	<b>Lyon en janvier.</b> Indice de sensibilité principal sur chaque orientation pour les gains par les fenêtres de chaque zone . . . . .	97

4.9	<b>Strasbourg en janvier.</b> Dispersion des 2000 valeurs de besoins de chauffage	99
4.10	<b>Strasbourg en janvier.</b> Dispersion sous forme de boxplot des pertes et gains par vitrage pour chaque zone et orientation . . . . .	100
4.11	<b>Strasbourg en janvier.</b> Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total . . . . .	102
4.12	<b>Strasbourg en janvier.</b> Indice de sensibilité principal sur chaque orientations pour les pertes par les fenêtres de chaque zone . . . . .	102
4.13	<b>Strasbourg en janvier.</b> Indice de sensibilité principal sur chaque orientations pour les gains par les fenêtres de chaque zone . . . . .	103
4.14	<b>Lyon en juillet.</b> Dispersion des 2000 valeurs de besoins de refroidissement	105
4.15	<b>Lyon en juillet.</b> Dispersion sous forme de boxplot des pertes et gains par vitrage pour chaque zone et orientation. . . . .	106
4.16	<b>Lyon en juillet</b> Indice de sensibilité principal sur le besoin de refroidissement des deux zones et le besoin total . . . . .	107
4.17	<b>Lyon en juillet.</b> Indice de sensibilité principal sur chaque orientation pour les pertes par les fenêtres de chaque zone . . . . .	108
4.18	<b>Lyon en juillet.</b> Indice de sensibilité principal sur chaque orientations pour les gains par les fenêtres de chaque zone . . . . .	108
4.19	<b>Strasbourg en juillet</b> Indice de sensibilité principal sur le besoin de refroidissement des deux zones et le besoin total . . . . .	110
4.20	<b>Strasbourg en juillet.</b> Indice de sensibilité principal sur chaque orientations pour les pertes par les fenêtres de chaque zone . . . . .	111
4.21	<b>Strasbourg en juillet.</b> Indice de sensibilité principal sur chaque orientations pour les gains par les fenêtres de chaque zone . . . . .	111
4.22	<b>Triple Vitrage en janvier</b> Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total . . . . .	114
4.23	<b>Triple Vitrage en juillet</b> Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total . . . . .	116
4.24	<b>Maison Dégradée en janvier</b> Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total . . . . .	119
4.25	<b>Maison Dégradée en juillet</b> Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total . . . . .	121
4.26	<b>Cas test original en janvier</b> Évolution des indices de sensibilité en fonction du besoin de chauffage total horaire . . . . .	124
4.27	<b>Maison Dégradée en janvier</b> Évolution des indices de sensibilité en fonction du besoin de chauffage total horaire . . . . .	124
4.28	<b>Triple Vitrage en janvier</b> Évolution des indices de sensibilité en fonction du besoin de chauffage total horaire . . . . .	125

4.29 <b>Cas test original en juillet</b> Évolution des indices de sensibilité en fonction du besoin de refroidissement total horaire . . . . .	125
4.30 <b>Maison dégradée en juillet</b> Évolution des indices de sensibilité en fonction du besoin de refroidissement total horaire . . . . .	126
4.31 <b>Triple Vitrage en juillet</b> Évolution des indices de sensibilité en fonction du besoin de refroidissement total horaire . . . . .	126

# Nomenclature

## Lettres majuscules

---

Notation	Unité	Signification
$C_p$	$[J.kg^{-1}.K^{-1}]$	Capacité calorifique
$Ens$	[h]	Durée d'ensoleillement
$H_R$	[%]	Humidité relative
$M_i$	[-]	Nombre de mode ou terme pour KL
$N_t$	[-]	Nombre de pas de temps horaire
$N_T$	[-]	Nombre d'évaluation du modèle
$N$	[-]	Nombre de tir aléatoire
Pr	[-]	Probabilité ; seuil de confiance
$R_{Dir}$	$[W.m^{-2}]$	Rayonnement solaire direct normal
$R_{Diff}$	$[W.m^{-2}]$	Rayonnement solaire diffus horizontal
$S_i$	[-]	Indice de sensibilité de première ordre
$S_T$	[-]	Indice de sensibilité total
$T$	[°C]	Température sèche extérieure
$T_m$	[°C]	Moyenne mensuelle de la température quotidienne
$T_n$	[°C]	Moyenne mensuelle de la température minimale quotidienne
$T_x$	[°C]	Moyenne mensuelle de la température maximale quotidienne
$V_{Dir}$	[°]	Direction de vent
$V_{Vit}$	$[m.s^{-1}]$	Vitesse de vent
$X$	[-]	Entrées du modèle
$Y$	[-]	Sortie du modèle

---

## Abréviations

---

Notation	Signification
BBC	Bâtiment Basse Consommation
DM	Double-Mur
FAST	Fourier Amplitude Sensitivity Test
IWEC	International Weather Energy Calculations
KL	Karhunen-loève
LHS	Latin Hypercube Sampling
MF	Météo-France
OAT	One-At-the-Time
OMM	Organisation Mondiale de Météorologie
RDJ	Rez-De-Jardin
STD	Simulation Thermique Dynamique
TMY	Typical Meteorological Year
UTC	Universal Time Coordinated
VMC	Ventilation Mécanique Contrôlée

## Opérateurs et fonctions

---

Notation	Signification
$f(\cdot)$	Modèle
$h(\cdot)$	Fonction de distribution
$C(\cdot)$	Fonction d'auto-corrélation
$V(\cdot)$	Variance
$E(\cdot)$	Espérance mathématique
$\bar{\cdot}$	Moyenne
$\hat{\cdot}$	Estimateur Bootstrap
$\mathcal{U}$	Loi uniforme
$\mathcal{N}$	Loi normale

## Lettres minuscules

---

Notation	Signification	Unité
$b$	[-]	b-ième réplique bootstrap
$k$	[-]	Nombre de réplique bootstrap
$p$	[-]	Nombre de variables d'entrée
$t$	[h]	Variable relative au temps
$r$	[-]	Caractérise l'intervalle de confiance

## Lettres grecques

---

Notation	Unité	Signification
$\theta$	[-]	Caractère stochastique
$\rho$	[-]	Coefficient de corrélation de Pearson
$\rho$	[kg.m <sup>-3</sup> ]	Masse volumique
$\lambda$	[W.m <sup>-1</sup> .K <sup>-1</sup> ]	Conductivité
$\lambda_{ki}$	[-]	Valeur propre
$\mu$	[-]	Moyenne
$\sigma$	[-]	Écart type
$\varepsilon$	[-]	Champ aléatoire
$\xi_{ki}$	[-]	Variable aléatoire

## Indices

---

Notation	Signification
$i_{wec}$	Relatif aux données du IWECC
$MF$	Relatif aux données de MF
$MoyMF$	Relatif à la moyenne des données de MF
$ech$	Relatif aux échantillons
$Boot$	Relatif au bootstrap



# Introduction

Ce travail de thèse s'inscrit dans le cadre du projet ANR FIABILITE qui porte sur la fiabilité des logiciels de simulation thermique dynamique et plus particulièrement sur les sources potentielles de biais et d'incertitude dans le domaine de la modélisation thermique et énergétique des bâtiments basse consommation l'objectif étant de déterminer ces sources et de quantifier les incertitudes qu'elles engendrent. On y étudie les hypothèses et approximations des modèles de simulation mais également les facteurs d'entrée nécessaires à la modélisation et qui décrivent l'objet bâtiment et son environnement. Ces facteurs peuvent être des entrées statiques qui restent constantes durant la simulation, comme la composition du bâtiment, ou dynamiques qui varient au cours de la simulation comme les données climatiques ou l'occupation. Ces outils de simulation sont des aides à la conception et les résultats de simulation conditionnent des choix importants dans la composition et le fonctionnement du bâtiment. Il est donc primordial d'obtenir une qualité de résultat de simulation suffisante afin que les performances du bâtiment basse consommation évaluées en phase conception puissent être à la hauteur des consommations réelles.

Une des missions du projet FIABILITE est d'accéder à cette qualité en développant de nouvelles méthodologies pour évaluer la dispersion sur les résultats de simulation et d'en extraire les entrées incertaines qui en sont responsables ce qui peut amener des pistes d'amélioration des modèles sous-jacents, une réduction des incertitudes des paramètres les plus influents si cela est possible ou tout simplement une prise de décision différente en fonction de la sensibilité de la solution aux incertitudes. Cette quête de fiabilité doit prendre en compte les incertitudes liées aux paramètres de conception mais aussi aux sollicitations et aux usages du bâtiment. Or, les sollicitations comme les données climatiques et les profils d'occupation sont rarement prises en compte dans le bilan d'incertitude bien qu'il s'agisse des variables essentielles pour accéder à la consommation d'un bâtiment. Cette lacune s'explique d'une part par la difficulté à quantifier la variabilité associée à ces entrées temporelles et d'autre part par le verrou scientifique que constitue leurs gestions dans les méthodes standard d'évaluation des paramètres influents sur la dispersion.

La motivation de ce travail de thèse est d'apporter une méthodologie de gestion des

incertitudes et de la sensibilité en simulation thermique dynamique lorsque les entrées incertaines sont temporelles et corrélées entre elles.

Le premier chapitre expose les enjeux de la simulation thermique dynamique dans le secteur du bâtiment ainsi que les incertitudes qui y sont liées. L'impact de différentes sources d'incertitudes sur la réponse du modèle est estimé pour une maison basse consommation qui servira de cas test dans le cadre de ce travail de thèse.

Dans le second chapitre on présente les méthodes d'analyses d'incertitude et de sensibilité qui permettent un gain en fiabilité des simulations thermiques dynamiques en évaluant la dispersion des résultats et en déterminant les entrées qui y contribuent significativement avant d'évoquer le verrou que constitue la gestion des entrées fonctionnelles en analyse de sensibilité.

Le chapitre 3 sera consacré à la mise en place de la méthodologie pour la gestion des entrées fonctionnelles corrélées en analyse de sensibilité mais également pour la caractérisation de la variabilité des entrées. La méthode sera appliquée au cas des données météorologiques à l'aide des fichiers climatiques utilisés en simulation thermique dynamique. Les résultats de génération des entrées seront ensuite confrontés à des observations réelles de température et d'ensoleillement.

Le dernier chapitre regroupera un ensemble d'applications sur le cas test de la maison basse consommation. Cette phase permettra de vérifier par propagation dans un modèle thermique la méthodologie développée. Différents climats seront étudiés, puis différentes configurations technologiques et enfin une analyse de sensibilité regroupant entrées statiques et entrées temporelles sera réalisée. Pour chaque configuration on évaluera la dispersion des besoins énergétiques liée à la variabilité météorologique et on en déduira par analyse de sensibilité les variables d'entrée responsables.

# Chapitre 1

## Contexte de l'étude et définition de la problématique : Incertitudes et bâtiment à haute efficacité énergétique

---

Le but de ce chapitre est de montrer l'importance de la gestion des incertitudes dans les logiciels de simulation thermique dynamique qui servent à l'estimation des performances énergétiques de bâtiments basse consommation. Pour cela on présente tout d'abord les enjeux de la simulation dans le secteur du bâtiment. Puis on expose quelques notions en incertitude afin de présenter les différentes étapes de la simulation thermique dynamique dans lesquelles l'incertitude est présente. Enfin quelques exemples de l'impact de ces incertitudes sur un bâtiment test sont proposés et amènent à la conclusion que la simulation thermique dynamique sur des bâtiments basse consommation d'énergie doit prendre en compte les incertitudes liées aux sollicitations météorologiques.

---

## 1.1 De la gestion des incertitudes en simulation thermique dynamique vers la garantie de performance

### 1.1.1 Nécessité d'un gain en fiabilité

#### 1.1.1.1 La modification des postes de dépense énergétique

De nos jours la préoccupation de la conservation des ressources énergétiques nécessite une réflexion sur la diminution de nos consommations. Or, le bâtiment est un des secteurs les plus consommateurs d'énergie en France avec 44% de l'énergie finale<sup>1</sup>. Dans ce contexte des projets ambitieux de réduction de la consommation énergétique du parc immobilier français sont mis en place avec notamment le Grenelle de l'environnement ainsi que des réglementations thermiques imposant des seuils de performance énergétique.

L'optimisation des performances énergétiques modifie la réponse du bâtiment aux sollicitations et amène à une redéfinition des phénomènes intervenant sur les gains et pertes énergétiques. En effet, dans les bâtiments classiques, les déperditions à travers les parois dominent le reste des déperditions, ce qui permet d'obtenir en quasi-statique une estimation des besoins de chauffage représentative, notamment par l'utilisation des degrés-jours. Cette approximation pour le cas de bâtiments basse consommation n'est plus adaptée car l'isolation des murs, les fenêtres double voire triple vitrage ainsi qu'une meilleure gestion des infiltrations diminuent l'impact de la température extérieure sur les besoins.

Des phénomènes comme l'inertie sont plus visibles puisqu'ils sont intensifiés par les nouvelles caractéristiques du bâtiment comme l'isolation qui pour l'inertie contribue au stockage de l'énergie. Les capacités de stockage des bâtiments aidant, l'impact des apports solaires sur les besoins de chauffage augmentent. Des études comme les travaux de Chesné et al. [8] montrent même que certains bâtiments basse consommation d'énergie peuvent s'affranchir des besoins de chauffage grâce à une bonne gestion des apports solaires. On en déduit que les enjeux dans ces bâtiments sont donc de récupérer, de stocker et de restituer efficacement l'énergie solaire reçue afin de diminuer les besoins de chauffage. De même, les apports internes dus aux équipements et aux occupants contribuent de manière plus importante à l'augmentation de la température intérieure dans un bâtiment basse consommation que dans un bâtiment classique. Cette addition de phénomènes dont on ne peut plus négliger l'impact dans les bâtiments basse consommation tend à réduire la corrélation forte qui existe entre température extérieure et consommation énergétique. A ce propos, une étude empirique [41] montre la difficulté d'établir une corrélation nette entre les consommations réelles d'un échantillon mondial de bâtiments dits de basse consom-

---

<sup>1</sup>Chiffres clés Énergie 2012 du gouvernement

## **1.1. DE LA GESTION DES INCERTITUDES EN SIMULATION THERMIQUE DYNAMIQUE VERS LA GARANTIE DE PERFORMANCE**

---

mation énergétique et les degrés jours associés à chaque bâtiment.

Les postes de dépense et de gain énergétique doivent être redéfinis ainsi que l'impact des variables météorologiques lorsque l'on considère un bâtiment basse consommation. Une approche dynamique est primordiale pour évaluer les besoins énergétiques. Chaque élément du bâtiment peut influencer sur sa consommation ce qui rend la compréhension de la gestion énergétique du bâtiment moins évidente pour les concepteurs.

### **1.1.1.2 La simulation en réponse aux attentes des concepteurs**

Pour concevoir des bâtiments basse consommation, on a recours à des outils de simulation numériques qui évaluent la performance énergétique du futur bâtiment et son potentiel d'amélioration. Certains de ces outils permettent de réaliser des simulations thermiques dynamiques, c'est-à-dire de reproduire sur l'année le comportement du bâtiment sous une certaine sollicitation météorologique en prenant en compte les caractéristiques du bâtiment étudié. Ces simulations tentent de représenter le comportement du bâtiment en fonctionnement réel c'est-à-dire avec des occupants (scénario, charge interne, etc.) et des usages (électroménager, luminaire, etc.). Les résultats de ces simulations permettent d'estimer la performance énergétique du bâti et d'évaluer les différents choix de solutions constructives.

Dans ce contexte de réduction globale des consommations énergétiques, se pose la question de garantir les performances d'un bâtiment de sa phase conception à sa phase d'exploitation. Les outils numériques de conception, qui pour certains ont été développés par des équipes de recherche afin de mieux comprendre les interactions au sein d'un système physique qu'est le bâtiment, sont désormais utilisés à des fins de labélisation et de réglementations. On tend donc à associer une valeur juridique aux résultats de simulation.

Cette garantie ne peut se faire sans prise en compte d'une marge d'erreur ou d'un intervalle d'incertitude. La performance énergétique d'un bâtiment évaluée par la Simulation Thermique Dynamique (STD) ne doit plus être une estimation de la performance mais plutôt une probabilité d'obtenir cette valeur en fonction des incertitudes composant la STD. Il devient donc primordial de gérer ces incertitudes de manière globale dans la simulation. Cette gestion passe par l'identification et l'évaluation de ces incertitudes.

La partie suivante traite de la définition des concepts et de la terminologie des incertitudes, puis suit leur identification à chaque étape de la STD.

## 1.1.2 Les incertitudes : concept et terminologie

Le but de cette partie est de mieux appréhender les termes et concepts liés aux incertitudes afin de comprendre les enjeux des chapitres suivants.

### 1.1.2.1 Comment définir une grandeur physique ?

Pour la compréhension de phénomènes et pour permettre la modélisation de la réalité, les scientifiques tentent de connaître la valeur d'une grandeur physique. Pour l'obtenir nous devons passer par l'observation de cette grandeur ou du phénomène lié à celle-ci par sa mesure. Cependant en répétant la mesure d'une grandeur avec le même protocole, on peut s'apercevoir que l'on obtient des résultats différents de la première mesure. Ces résultats sont certes proches les uns des autres mais différents. Dans la plupart des cas on reste proche d'une valeur moyenne mais il se peut que des résultats en soient plus éloignés ce qui correspond à une dispersion.

Une des raisons de cette dispersion est la variation des conditions de l'expérience et la variabilité naturelle de la grandeur. Par exemple, si on mesure la vitesse de propagation d'un son, entre deux mesures, les caractéristiques physiques de l'air (température, pression, humidité, etc.) auront légèrement variées, suffisamment pour entraîner une modification de la valeur mesurée. Ces variations (des conditions extérieures et par conséquent de la grandeur mesurée) sont inévitables. Selon Konstantin Protassov [51], "nous sommes "condamnés" à effectuer des mesures de grandeurs qui ne sont presque jamais constantes". Ainsi une grandeur se caractérise non pas par une valeur mais par la probabilité d'obtenir telle ou telle valeur. Pour la grande majorité des grandeurs physiques cette probabilité se traduit par une distribution de Gauss (ou loi normale). Dans ce cas, la grandeur physique est définie par la valeur moyenne de la distribution associée à sa dispersion.

### 1.1.2.2 Quelques notions sur l'incertitude

#### **Incertitude aléatoire et incertitude systématique**

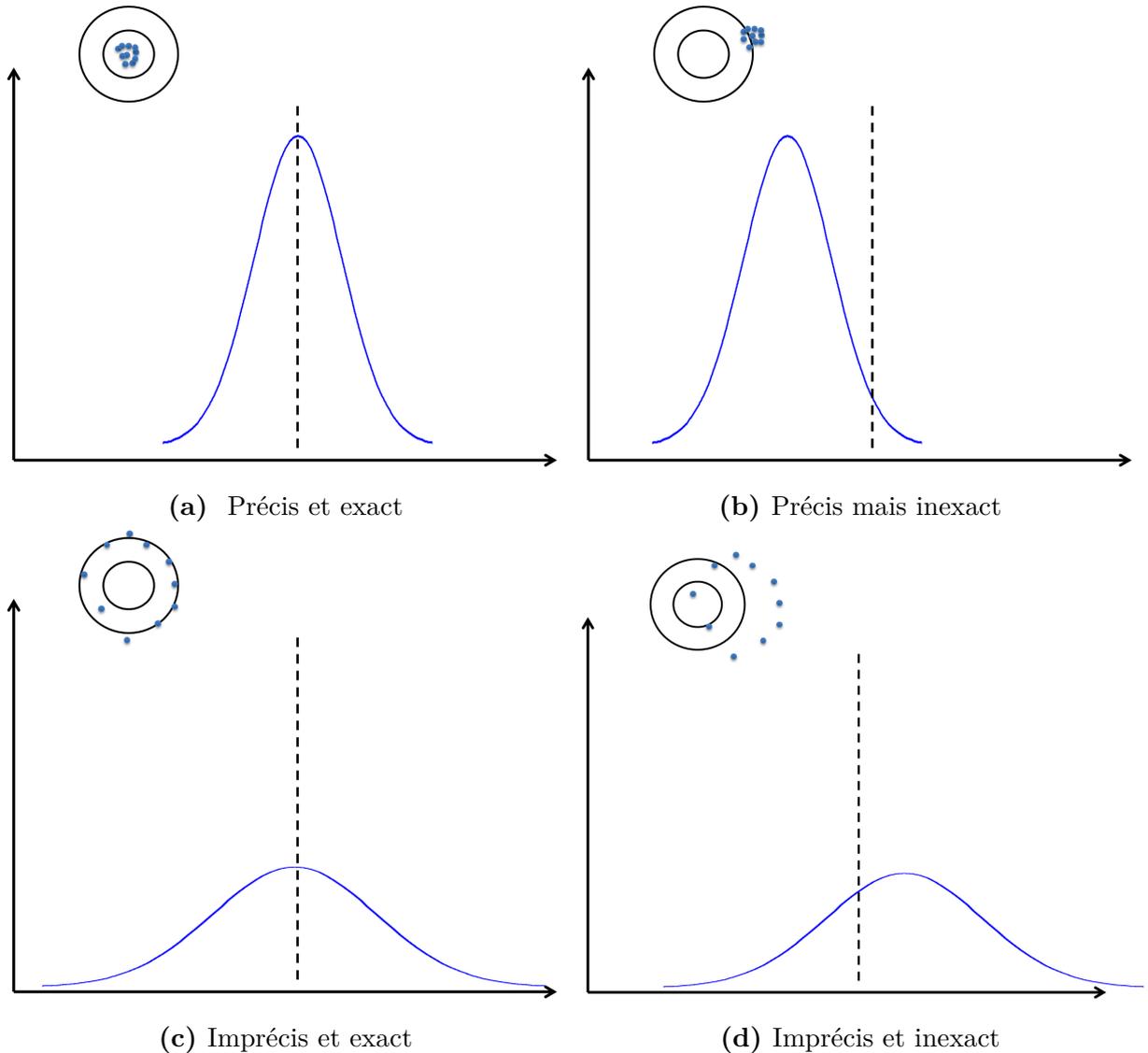
L'incertitude systématique est un synonyme de biais, il s'agit de l'écart entre le résultat (la valeur mesurée) et la valeur vraie. Plus le biais est important plus on est éloigné de la valeur réelle. Ceci est dû à la méconnaissance du système, au fait que les modèles ne sont qu'une représentation de la réalité, ce qui, par conséquent implique une erreur systématique. L'incertitude aléatoire quant à elle, est la variable aléatoire représentant la variabilité naturelle des grandeurs physiques. Elle est inversement proportionnelle à la précision, il s'agit de la dispersion.

#### **Exactitude et précision**

Un résultat exact ne présente pas de biais, l'écart à la valeur réelle est nul. Un résultat

## 1.1. DE LA GESTION DES INCERTITUDES EN SIMULATION THERMIQUE DYNAMIQUE VERS LA GARANTIE DE PERFORMANCE

précis présente une faible incertitude aléatoire. La dispersion des valeurs est faible. Pour imaginer ces deux concepts nous allons utiliser l'analogie connue dans la pédagogie de l'incertitude du tireur d'élite sur une cible. Sur la figure 1.1 à chaque analogie est associée le visuel de la distribution de Gauss que l'on pourrait observer.



**Figure 1.1** – Illustration des différentes configurations de précision et d'exactitude par analogie avec une cible et le type de gaussienne correspondante. La cible est représentée en haut à gauche de chaque sous-figure. La ligne en pointillés est la valeur du centre de la cible.

Sur la figure 1.1a, en haut à gauche est représentée la cible, on observe que le tireur d'élite a visé juste, et l'écart entre ses tirs est petit, le tir est exact et précis. La gaussienne associée est donc peu dispersée et sa moyenne correspond à la valeur de référence (en

pointillés). En figure 1.1b le tir est précis mais inexact, il y a peu d'écart entre chaque tir mais ce tir est à chaque fois en dehors de la cible. La gaussienne représentée est étroite mais décalée par rapport à la valeur de référence. Le biais est grand mais la dispersion faible. Dans le cas de la figure 1.1c et 1.1d la dispersion est grande, la gaussienne étalée, le tir est peu précis. Dans le premier cas, contrairement au second, les tirs englobent la valeur exacte (de référence). Tout scientifique souhaite un résultat exact et précis.

Dans le cadre de la modélisation, pour quantifier le biais et connaître l'exactitude d'un modèle, il faut confronter le modèle à la réalité, il s'agit d'une étape de la validation de modèle. Pour cela des valeurs expérimentales sont comparées aux valeurs numériques. Une des phases du projet FIABILITE dans lequel s'inscrit ce travail prévoit la validation de modèle de STD grâce à des cas tests simples appliqués à des bâtiments expérimentaux dont disposent les partenaires du projet (plateforme PREDIS, cellules Bestlab, Maisons INCAS). La validation de modèle est un processus délicat avec une méthodologie bien définie (voir [46], [47], [28], [31] pour plus d'informations).

Le thème développé dans ce manuscrit ne concerne pas la validation de modèle mais l'évaluation de la précision de la simulation en déterminant la dispersion sur les résultats de simulation. On peut en effet améliorer les modèles et réduire le biais afin d'obtenir un résultat proche de l'exactitude, mais si ce résultat est peu précis (grand intervalle d'incertitude) il devient tout aussi inutilisable dans le contexte des garanties de performance dans les bâtiment dits de basse consommation.

### Notion d'intervalle et de niveau de confiance

Lorsque la distribution de la grandeur est supposée gaussienne ( $\mathcal{N}(\mu, \sigma)$ ), l'intervalle de confiance se note  $\bar{x} \pm r\sigma$  avec  $r$  le coefficient multiplicatif qui traduit la taille de l'intervalle de confiance. On prend la moyenne de la grandeur  $x$  pour la valeur "réelle" et sa largeur  $\sigma$  pour l'erreur. Cette écriture révèle une interprétation en termes de probabilité qu'il est important de bien visualiser. La probabilité  $\Pr$  de trouver la valeur  $x$  dans l'intervalle  $[\mu - \sigma, \mu + \sigma]$  est de 68,27% telle que :

$$\Pr[\mu - \sigma, \mu + \sigma] \simeq 68,27\% \tag{1.1}$$

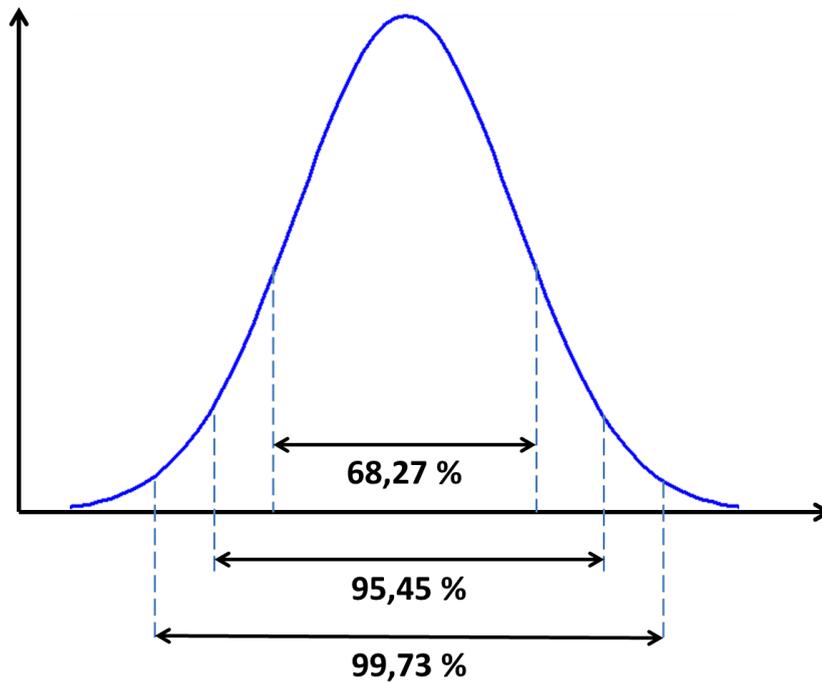
si  $x$  est compris dans l'intervalle  $[\mu - 2\sigma, \mu + 2\sigma]$  on a :

$$\Pr[\mu - 2\sigma, \mu + 2\sigma] \simeq 95,45\% \tag{1.2}$$

dans l'intervalle  $[\mu - 3\sigma, \mu + 3\sigma]$  :

$$\Pr[\mu - 3\sigma, \mu + 3\sigma] \simeq 99,73\% \tag{1.3}$$

## 1.1. DE LA GESTION DES INCERTITUDES EN SIMULATION THERMIQUE DYNAMIQUE VERS LA GARANTIE DE PERFORMANCE



**Figure 1.2** – Représentation sur la distribution de gauss des probabilités de trouver  $x$  dans un intervalle de  $2\sigma$  (68.27%),  $4\sigma$  (95,45%) et  $6\sigma$  (99,73%)

La figure 1.2 illustre la part de la distribution de gauss associée à la probabilité de trouver la valeur  $x$  dans l'intervalle considéré. Cette probabilité est le niveau de confiance que l'on choisit en fonction de l'intervalle de confiance considéré. Dans la suite de ce travail, lorsque la notation du type  $x = \mu \pm r\sigma$  est utilisée, le niveau de confiance est renseigné (le seuil de probabilité). L'équivalence entre intervalle de confiance et les niveaux de confiance est résumée dans le tableau 1.1.

**Table 1.1** – Probabilité Pr (en %) pour que la valeur d'une variable gaussienne  $x$  soit dans l'intervalle  $[\mu - r\sigma, \mu + r\sigma]$  pour diverses valeurs de  $r$

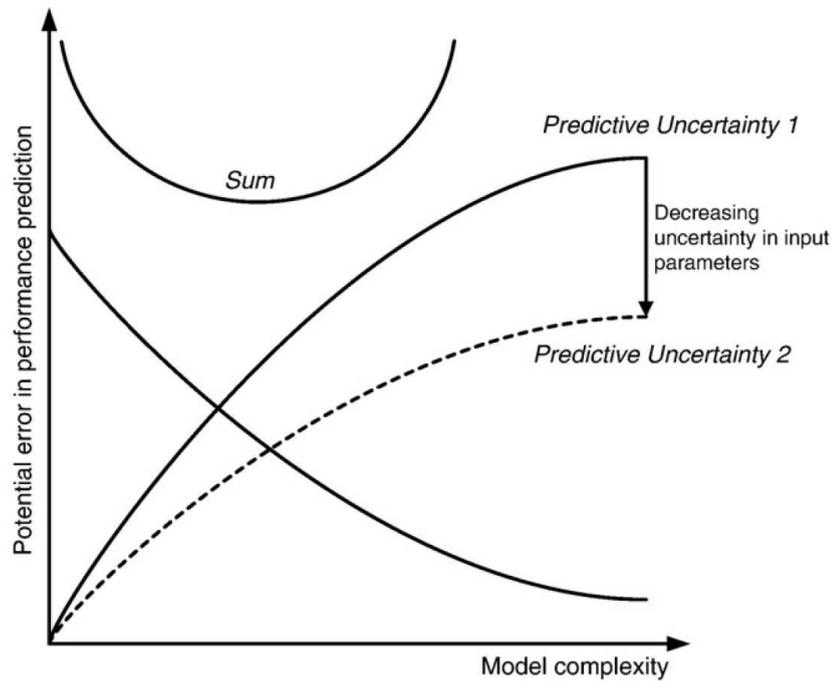
Pr	95,0	99,0	99,9
r	1,960	2,576	3,290

### 1.1.3 Discussion sur la présence des incertitudes en simulation thermique dynamique

#### 1.1.3.1 Vers des logiciels de plus en plus complexes

La communauté internationale de la thermique du bâtiment développe et valide des modèles utilisés dans la STD depuis plus de 40 ans [24]. Ces modèles font intervenir un grand nombre de domaines : l'aérodynamique, la météorologie, l'hygro-thermique, parfois même la sociologie et la biologie avec les modèles de confort et le comportement des occupants, etc. Cette interdisciplinarité rend les modèles utilisés dans la STD, certes de plus en plus proches de la réalité (donc diminution du biais), mais aussi de plus en plus complexes. La complexité du modèle détermine le nombre de paramètres d'entrées car plus un modèle devient complexe plus la quantité d'informations dont il a besoin augmente. On va donc avoir besoin d'augmenter la connaissance du système pour répondre à cette complexité. Or la méconnaissance du système est une problématique récurrente dans le domaine du bâtiment. Certains aspects sont peu maîtrisés et notamment les interactions entre les différents domaines (ex : impact du métabolisme humain sur l'aérodynamique d'une pièce). Et cette méconnaissance du système ajoutée à l'augmentation de la complexité des modèles provoque une augmentation de l'incertitude sur les sorties de la simulation thermique dynamique comme l'illustre le graphe en figure 1.3 extrait des travaux de Marija Trčka et al.[67].

## 1.1. DE LA GESTION DES INCERTITUDES EN SIMULATION THERMIQUE DYNAMIQUE VERS LA GARANTIE DE PERFORMANCE



**Figure 1.3** – Représentation de l’incertitude potentielle dans les prédictions de performance en STD en fonction de la complexité du modèle [67]

La gestion des incertitudes dans les modèles complexes est une problématique que tente de gérer la communauté de la simulation en général. Un exemple de projets sur cette problématique est le programme de recherche MUCM<sup>2</sup> (Managing Uncertainty in Complex Models) commencé en 2006 et avec de nombreux domaines d’application (aéronautique, astrophysique, biologie, etc.). De même le GdR MASCOT-NUM<sup>3</sup> (Méthodes d’Analyse Stochastique pour les COdes et Traitements NUMériques) est une collaboration française travaillant sur ces problématiques. Par ailleurs dans le domaine de la simulation du bâtiment, il existe des travaux de recherches traitant de la gestion de ces incertitudes et notamment par les techniques d’analyse de sensibilité, nous y reviendrons plus loin dans le manuscrit, puisqu’il s’agit du type de méthode exploité dans ce travail de thèse.

La problématique concerne la confiance dans les résultats de simulation et on cherche à obtenir les plages d’incertitudes sur des variables de sortie. Or, les sources d’incertitudes dans le processus de simulation thermique dynamique (STD) peuvent être nombreuses. La prochaine partie tente de les présenter en fonction des différentes étapes de la STD.

<sup>2</sup>[www.mucm.ac.uk](http://www.mucm.ac.uk)

<sup>3</sup>[www.gdr-mascotnum.fr](http://www.gdr-mascotnum.fr)

### 1.1.3.2 Les incertitudes aux différentes étapes de la simulation thermique dynamique

La STD peut être décomposée de manière schématique (figure 1.4) par un noyau de calcul avec différents modèles dont les entrées doivent être renseignées. Ces entrées peuvent être de différentes natures : statique (dit paramètre), dynamique (dit sollicitation), ou un choix entre différents sous-modèles pour obtenir une entrée (hypothèses de simulation) comme par exemple le choix de la corrélation pour le calcul des coefficients de convection. On obtient les sorties de simulation, ce qui implique ensuite l'interprétation de ces résultats qui peuvent être, eux aussi, statiques (ex : la consommation annuelle du bâtiment) ou dynamiques (ex : évolution de la température intérieure).

L'utilisateur logiciel (à ne pas confondre avec l'utilisateur du bâtiment) est présent tout le long du processus de simulation. Il en est le gestionnaire, autant dans les choix qu'il fait, les erreurs et les interprétations. C'est la raison pour laquelle en figure 1.4 il englobe schématiquement tout le processus de simulation.

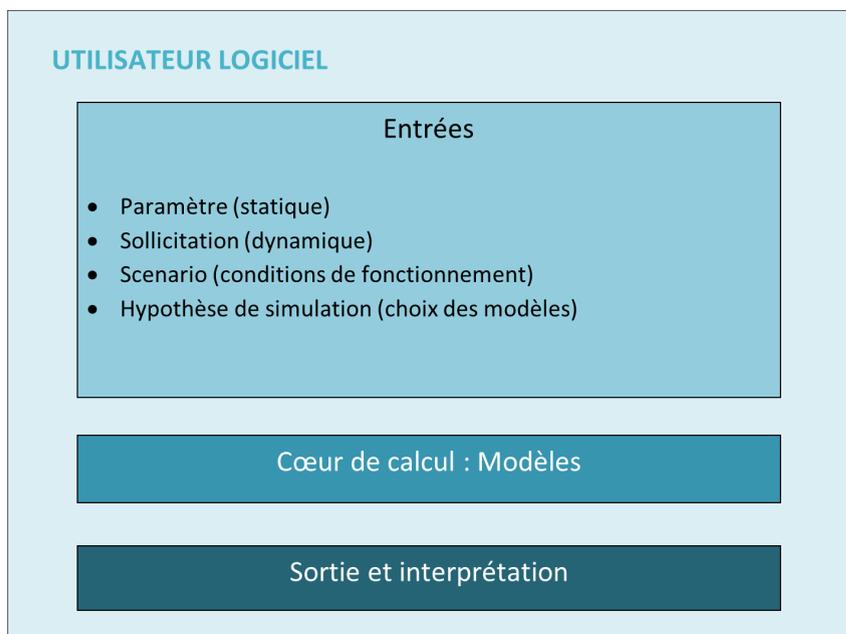


Figure 1.4 – Décomposition schématique de la simulation thermique dynamique

#### Les données d'entrée

Chaque grandeur physique a une incertitude associée à sa valeur. Il y a certes la variabilité naturelle de la grandeur mais la plus grande part réside dans la méconnaissance même de cette valeur. Il faut donc considérer ces incertitudes et les propager dans le modèle

## **1.1. DE LA GESTION DES INCERTITUDES EN SIMULATION THERMIQUE DYNAMIQUE VERS LA GARANTIE DE PERFORMANCE**

pour obtenir un résultat représentatif permettant une interprétation. Face à un système complexe faisant intervenir un grand nombre de modèles et par conséquent d'entrées imprécises, il devient difficile pour l'utilisateur de lier sorties et entrées. De plus, ces liens dépendent du modèle de bâtiment, de son mode de fonctionnement et des conditions limites (météorologie, occupant, etc.). Ainsi cette complexité masque la dépendance de la sortie à des hypothèses de simulation fortes et mène ainsi à une mauvaise interprétation donc à des conclusions erronées.

Pour pouvoir interpréter, on doit disposer en sortie d'une trace de l'impact des incertitudes d'entrées : des intervalles d'erreur sur les variables d'intérêt. Sans cette considération de la sortie comme une plage de variation et non habituellement comme une valeur fixe, l'interprétation est fautive puisque basée sur un résultat déterministe, les aspects probabilistes en entrée étant négligés à tort.

### **Le modèle**

La présence d'un biais entre les résultats de simulation et la réalité est inévitable, du fait que la simulation n'est qu'une représentation de la réalité. Des hypothèses de simulation et des simplifications doivent être faites. En complexifiant les modèles on tente de limiter le biais au niveau de la représentation de la réalité mais on l'augmente dans d'autres aspects de la simulation comme l'augmentation des incertitudes de sortie mais également l'augmentation d'erreurs de l'utilisateur.

### **La sortie et son interprétation**

La variable d'intérêt observée en STD est le plus souvent la consommation d'énergie d'un bâtiment. Mais le choix même de cette sortie est déjà l'introduction d'un biais, car une seule valeur résume une multitude de phénomènes physiques en interaction. Il faut être prudent car il y a des hypothèses sous-jacentes. L'agrégation de données et donc l'obtention d'un unique scalaire pour une simulation annuelle complexe limite l'interprétation, la transparence du comportement du bâtiment et la détection de possibles erreurs de simulation. On doit prendre en compte les incertitudes englobant les données d'entrées afin d'obtenir des intervalles de confiance en sortie et permettre une interprétation juste des performances d'un bâtiment. On doit aussi identifier les entrées les plus influentes.

### **L'utilisateur**

L'utilisateur logiciel est une des plus grandes sources d'erreur dans le processus de la simulation comme le montre Guyon [20] où plusieurs participants réalisent la simulation d'un même bâtiment. Les résultats montrent plus de 40% de dispersion par rapport à la valeur moyenne des valeurs de consommation calculées. Plusieurs catégories d'utilisateur ont participé à l'étude : des utilisateurs réguliers et d'autres moins. L'étude a consisté

ensuite à comparer les résultats des participants à la simulation de référence pour tester la dispersion des simulations et de visualiser le biais induit par l'utilisateur. Cette comparaison entre simulation de référence (très proche des valeurs mesurées dans les conditions de validation des modèles) et simulation des participants montre que ces derniers ont tendance à surestimer les consommations.

Si on transpose le concept des incertitudes à cette étude, on s'aperçoit que l'écart entre la valeur moyenne et la valeur de référence est le biais induit par l'utilisateur et que par contre la dispersion autour de la moyenne est l'incertitude induite par le changement d'utilisateur. Avec le même type d'exercice, une étude du BRE (Building Research Establishment, U.K.) [4] montre une variation de la valeur des consommations obtenues de -46% à +106% par rapport à la valeur de référence. Ces écarts, évidemment, dépendent de la complexité du bâtiment et de son utilisation. Une maison indépendante dans le cadre de l'étude française est plus simple à simuler que les bâtiments tertiaires modélisés par les participants de la seconde étude.

L'auteur de l'étude française ([20]) conseille aux utilisateurs de STD de faire vérifier leur simulation par des tiers. L'échange des compétences semble être une bonne solution pour limiter les erreurs des utilisateurs puisque la STD devient de plus en plus multidisciplinaire et qu'il n'est pas forcément évident de maîtriser tous les aspects des phénomènes interagissant dans un bâtiment. Par ailleurs en fonction du logiciel de STD utilisé, les utilisateurs peuvent induire des biais en fonction des affinités de chacun et de leur formation. Un programme de limitation des erreurs en STD rend possible de tester sa propre capacité à induire des biais en fonction de l'aspect étudié. Ce programme, BESTEST (Building Energy Simulation Test) est à la base une méthode pour tester, diagnostiquer et valider les logiciels de simulations avec des cas tests simples. Lors de la prise en main d'un logiciel qui a déjà été validé, cette méthode peut être utilisée pour tester les capacités de l'utilisateur à induire des erreurs dans ses simulations et identifier rapidement leur provenance. Les faiblesses de l'utilisateur sont ciblées et cela peut l'amener à être plus vigilant.

La complexité des modèles peut aussi induire plus d'erreurs de la part des utilisateurs par rapport à des modèles simples. La réduction d'un biais en approximant mieux la réalité risque d'en créer à cause de l'utilisateur et de l'augmentation de données d'entrée qui peuvent être imprécises. Ainsi, dans le cadre du logiciel CLIM2000 testé dans [20], l'auteur décide, face au grand nombre de participants faisant des erreurs sur le modèle de pont thermique, d'opter pour un modèle réduit.

Ce contexte montre la nécessité pour les utilisateurs de STD de contrôler leur simulation et de regagner de la transparence. C'est dans cet objectif que doit se développer des

outils d'aide à la décision qui doivent permettre d'identifier les entrées influentes sur la sortie considérée et de pouvoir visualiser de manière plus synthétique les différents phénomènes en jeu et leur interaction.

La partie suivante illustre l'impact associé à deux types d'incertitudes sur une maison basse consommation d'énergie : une entrée statique et une hypothèse de simulation.

## **1.2 Exemple de sources d'incertitudes et de leurs impacts sur une maison basse consommation**

La modification des performances énergétiques des bâtiments amène à reconsidérer les modèles pris en compte et l'impact de certaines entrées. Des aspects qui pouvaient être négligeables le sont moins. La partie ci-dessous présente tout d'abord le bâtiment de stratégie passive qui va servir de cas d'étude. Puis, quelques exemples sont présentés pour illustrer l'influence non négligeable de la méconnaissance d'une entrée, ou d'un choix d'hypothèse de simulation.

### **1.2.1 Présentation du cas d'étude : Maison INCAS DM**

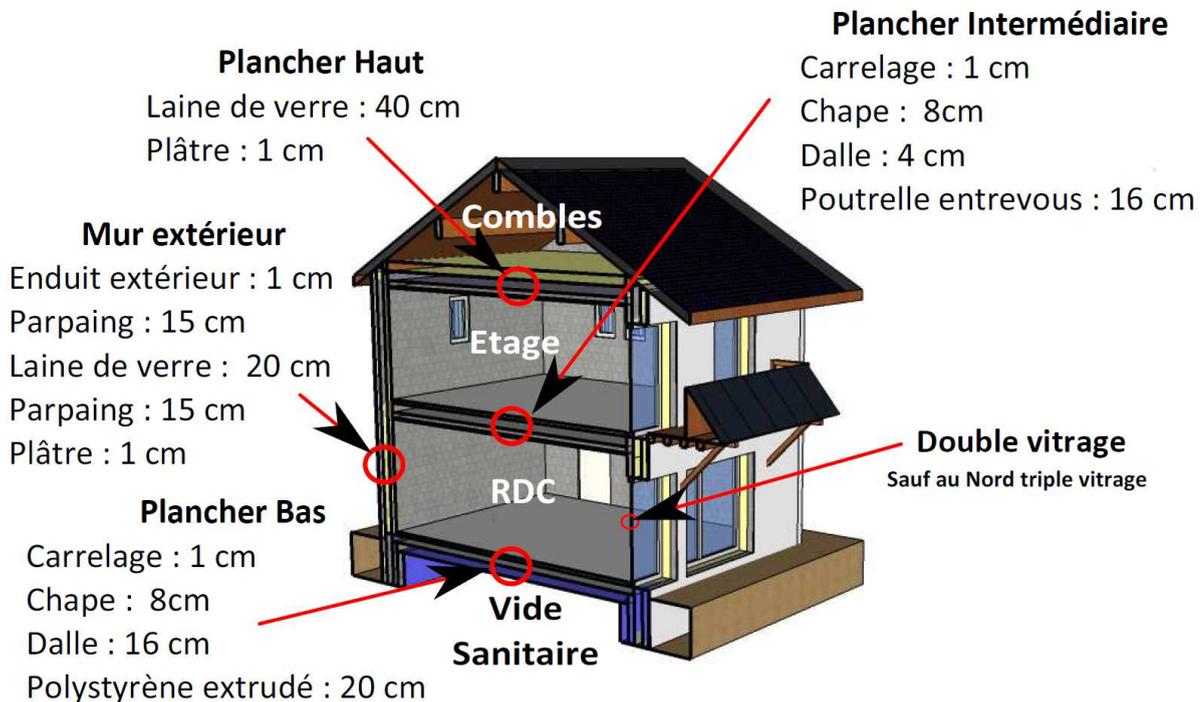
Le cas d'étude présenté ici reste le même dans les différentes parties de cette thèse et plusieurs études numériques sont réalisées. Il s'agit d'un des bâtiments tests exploités dans le projet ANR FIABILITE du programme "bâtiments et villes durables".

#### **1.2.1.1 Description du bâtiment test**

Le bâtiment test fait partie d'une plateforme expérimentale : la plateforme INCAS basée sur le site de l'INES (Institut National d'Énergie Solaire) au Bourget-du-Lac en Rhône-Alpes. Un des objectifs de cette plateforme est la validation de modèle. Elle est composée de plusieurs bâtiments basse consommation, de la même géométrie simple et compacte, mais avec des techniques constructives et des équipements différents. Chaque construction est instrumentée de plus d'une centaine de capteurs.

La maison étudiée est la Double-Mur (DM). Elle doit son nom à sa technique constructive qui consiste à placer l'isolation entre deux murs. On a ainsi, 15 cm de parpaing, 20 cm d'isolant et de nouveau 15 cm de parpaing. Les murs ont une épaisseur de 50 cm et la composition détaillée des parois est fournie en annexe. Un schéma résumant les principaux matériaux de la maison DM est présenté en figure 1.5. La technique constructive confère à la maison une certaine inertie qui peut se caractériser par le déphasage entre la sollicitation appliquée au bâtiment et sa réponse, ce qui permet un stockage de l'énergie

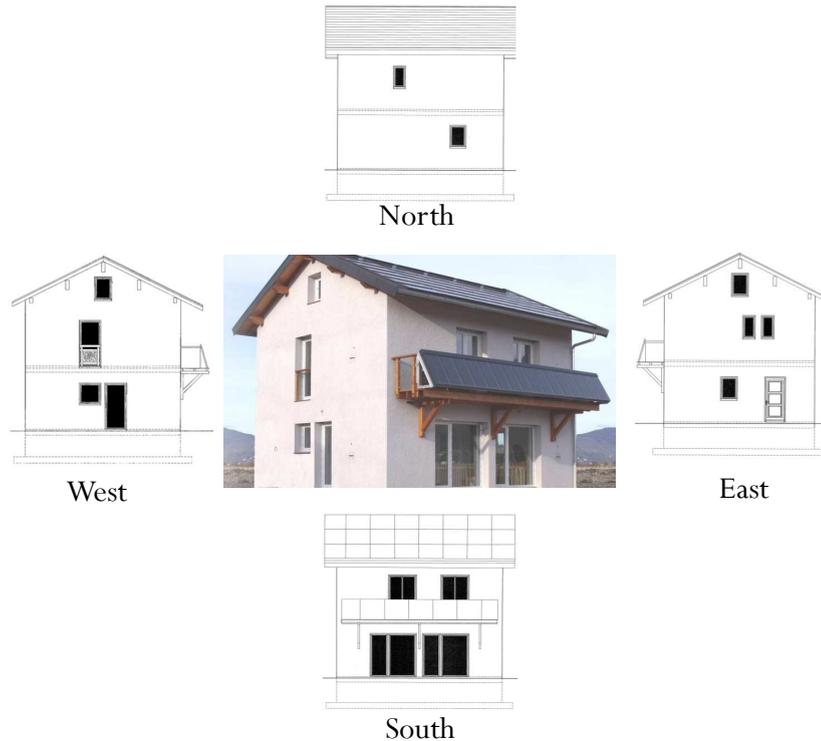
reçue par le bâtiment et une restitution quand cela est nécessaire. Par exemple, en hiver, elle permet d'emmagasiner les apports gratuits issus du rayonnement solaire et contribue à les conserver plus longtemps réduisant ainsi les consommations énergétiques liées aux chauffages à condition qu'elle soit accompagnée d'une isolation performante.



**Figure 1.5** – Principaux matériaux composant l'enveloppe du cas test INCAS-DM [61]

La maison dispose de deux grandes baies vitrées en double vitrage au sud pour maximiser les gains solaires en hiver. Ces gains sont limités l'été à l'aide de la fonction de casquette du toit et du balcon. On observe cette disposition des vitrages en fonction de l'orientation sur la figure 1.6. Côté nord les pertes thermiques sont minimisées par de petites fenêtres en triple vitrage, tandis que les autres correspondent à du double vitrage.

Le bâtiment est sur 4 niveaux, le vide sanitaire, le rez-de-jardin, l'étage et les combles. Une forte isolation entre les combles et l'étage ainsi qu'entre le vide sanitaire et le rez-de-jardin limite les échanges thermiques avec le sol ou le toit. Les bâtiments ont été conçus pour avoir une faible perméabilité à l'air, le taux d'infiltration est de 0.6 vol/h à 50 Pa. Le but est de limiter l'infiltration source de déperditions thermiques parasites qui sont particulièrement gênantes car elles ne sont pas contrôlées. On limite ainsi l'incertitude sur cette grandeur et sur le phénomène d'infiltration de manière globale. Par ailleurs, une attention particulière a été prise pour éviter les ponts thermiques qui entraînent sur la surface intérieure de la paroi une chute locale de température et crée des zones froides



**Figure 1.6** – Prise de vue de la maison DM associée aux croquis de la distribution des vitrages en fonction des 4 orientations.

dans un bâtiment. Ces phénomènes sont complexes à modéliser.

La maison DM a une stratégie d'optimisation énergétique dite passive. Grâce à ses caractéristiques d'enveloppe elle tente d'optimiser les gains et pertes énergétiques. Le faible taux d'infiltration, l'isolation importante de la paroi et des vitrages limitent les pertes. La distribution des vitrages en fonction des orientations et le couplage à des casquettes permettent de minimiser et maximiser les gains solaires en fonction de la saison réduisant ainsi les consommations énergétiques liées au chauffage ou au refroidissement.

### 1.2.1.2 Description du modèle associé

#### Le logiciel de simulation EnergyPlus

EnergyPlus est un logiciel de simulation thermique dynamique à l'initiative de l'US Department of Energy qui est disponible depuis 2001 [12].

Le schéma en figure 1.7 est tiré de la documentation d'EnergyPlus [44] et représente les différents éléments du logiciel. La description du bâtiment et les résultats (en bleu foncé dans le schéma) sont des fichiers textes ASCII, ce qui, nous le verrons plus tard, va permettre une automatisation des simulations. Dans le cas présenté ici, la géométrie du bâtiment est simulée à l'aide d'une interface graphique (GoogleSketchup) mais les autres caractéristiques du bâtiment (composition des murs, environnement, etc.) sont renseignées directement dans le logiciel.

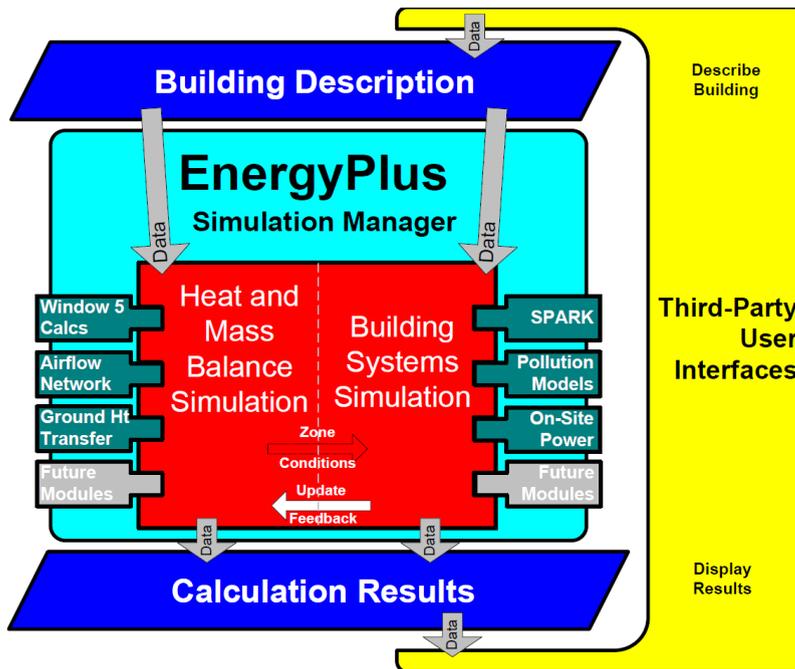


Figure 1.7 – Hiérarchie et besoin d'EnergyPlus

EnergyPlus dispose de nombreuses fonctionnalités et une base importante de modèles validée qui est disponible gratuitement [44]. Ces atouts en font un des logiciels de simulation thermique dynamique le plus utilisé et alimenté par la communauté internationale.

### Hypothèses de simulation

La maison est séparée en 4 zones thermiques : le rez-de-jardin, l'étage, le vide sanitaire et les combles. Les échanges aérauliques entre zones ne sont pas considérés. Une zone thermique est un volume où l'on considère la température comme homogène. Dans le cas de la simulation de la maison DM, afin de simplifier les calculs, aucun système n'est considéré : ni ventilation ni climatisation. Ainsi, on ne s'intéresse, pour la modélisation du bâtiment, qu'à l'aspect passif de la maison c'est-à-dire son enveloppe comprenant les murs, les planchers, les plafonds ainsi que les fenêtres. Les ombrages dus aux casquettes sont aussi considérés. Le bâtiment est en évolution libre.

Le milieu extérieur est pris en compte par le fichier météo de Chambéry mais aussi par différentes grandeurs imposées par l'utilisateur dans la modélisation telles que l'albédo, la température du sol, les masques, etc. Les expériences et les simulations se font sans gains internes (ni occupation, ni équipement pouvant dégager de la chaleur).

Des casquettes aux bords de fenêtres ont été rajoutées afin de simuler l'ombrage dû au retrait de la fenêtre par rapport à la façade extérieure du mur de 50 cm. Ces "casquettes virtuelles" sont d'une largeur de 35 cm.

## **1.2.2 Impact de l'incertitude de l'albédo**

### **1.2.2.1 Description de l'entrée**

L'albédo est la capacité de l'environnement à réfléchir le rayonnement solaire. Il se caractérise par le rapport entre le rayonnement réfléchi et le rayonnement solaire incident. Il s'agit d'un nombre adimensionnel compris entre 0 et 1. Plus le sol est clair et plus la part de rayonnement réfléchi est importante et plus l'albédo est grand.

### **1.2.2.2 Un paramètre incertain et variable**

La connaissance de l'albédo est assez incomplète puisque celui-ci varie en fonction de plusieurs variables telles que la nature du sol (herbe, sable, ciment, etc.) mais également des conditions météo (pluie, neige, etc.). Communément en milieu urbain l'albédo est supposé proche d'une valeur de 0,2. Or dans le cas des maisons INCAS le revêtement aux alentours est blanc soit un albédo plus proche de 0,4 (ciment clair). L'albédo varie suivant diverses composantes terrestres, d'après la littérature, il est supérieur à 0,8 dans le cas de la neige fraîche, et inférieur à 0,1 dans le cas des mers et océans. On retrouve dans le tableau 1.2 des valeurs d'albédo obtenues dans la littérature [2].

**Table 1.2** – Variabilité de l'albédo en fonction de la nature du sol

Nature du sol	Valeur de l'albédo
Béton	0,50
Herbe	0,2 à 0,3
Terre	0,15
Gravier	0,2
Neige fraîche	0,85

Cependant les valeurs peuvent varier en fonction des conditions de mesure, de la couleur et du type de revêtement. Le fait que cette entrée ne soit pas constante rend sa ca-

ractérisation plus difficile, ce qui peut expliquer les différences en fonction des références consultées. Une étude récente par Li et al. [35] évalue la variabilité de l'albédo en fonction de la journée, de la saison et du revêtement considéré. Dans les conditions de campagne de mesure, les auteurs établissent que la valeur de l'albédo ainsi que sa variabilité est importante en début et fin de journée. Pour obtenir une valeur plus basse et constante, la mesure doit être faite en pleine journée par ciel découvert. L'évaluation de l'albédo sur la surface terrestre est un sujet qui prend de l'ampleur notamment en physique de l'atmosphère où la valeur de l'albédo conditionne le réchauffement climatique. L'utilisation des données satellites permet une caractérisation de l'albédo de plus en plus fine spatialement.

Dans le cadre de la thermique du bâtiment, l'albédo est susceptible de provoquer un biais ou une incertitude car il joue directement sur la quantité de flux solaire réfléchi par le sol. Il impacte sur la température intérieure et donc sur les besoins énergétiques (chaud et froid). Dans le cas de maison à stratégie passive, l'aspect solaire impacte sur la température d'air intérieur. La mauvaise estimation ou la méconnaissance de ce paramètre peut entraîner un impact sur le comportement de la maison DM.

Une pré-étude expérimentale pour évaluer la variabilité de l'albédo in-situ a été faite par l'équipe du CEA-INES LEB sur le site de la plateforme. Sur la plage de temps du 15/04/2011 11 :00 au 27/04/2011 8 :00, en ne considérant que les données de 7h00 à 16h00, la valeur de l'albédo est présentée dans le tableau 1.3.

**Table 1.3** – Mesures préliminaires sur site de l'albédo

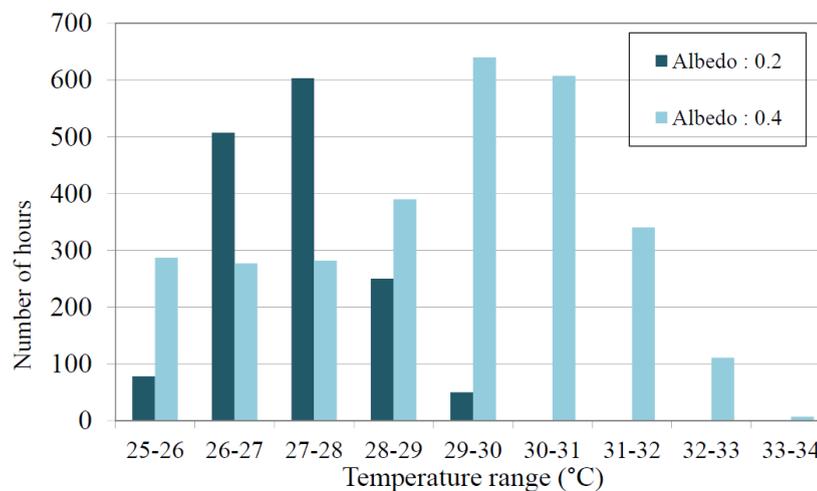
Type d'information	Valeur
Moyenne	0,39
Max	0,43
Min	0,25

Cette étude est une mesure approximative qui permet d'obtenir un ordre de grandeur de la valeur de l'albédo sur le site et de se rendre compte de sa variabilité sur des journées entières. On constate que sur une période relativement courte, soit moins de deux semaines l'albédo varie fortement. A noter que dans la simulation thermique dynamique par EnergyPlus cette entrée est considérée comme constante ce qui induit une source de biais. On peut toutefois choisir une valeur d'albédo pour chaque mois, mais la difficulté est alors de caractériser cette valeur en fonction du mois de l'année. Selon l'étude [35] aucune variabilité saisonnière n'a été observée, cependant il s'agit des conclusions obtenues sous les conditions expérimentales de l'étude (type de revêtement, climat californien, etc.). L'albédo est donc une entrée incertaine qui peut induire une incertitude et un biais

sur les résultats de simulation.

### 1.2.2.3 Résultats de l'impact sur la maison INCAS-DM

Pour évaluer l'impact de l'albédo sur les performances de la maison INCAS-DM, une étude numérique a été effectuée. Deux simulations ont été réalisées avec EnergyPlus, une avec un albédo de 0,2 (valeur par défaut dans les villes) et une avec un albédo de 0,4 (valeur plausible dans notre cas d'étude). On considère comme résultats de simulation l'évolution de la température opérative qui est plus proche de la notion de confort que la température intérieure. Afin d'obtenir un indicateur simple du confort d'été, on considère qu'au-delà de 27°C de température intérieure opérative, le bâtiment est en surchauffe et on décompte le nombre d'heure où la température opérative dépasse la valeur seuil. On obtient un équivalent de 40 jours de surchauffe pour un albédo de 0,2 contre 99 jours pour un albédo de 0,4. La figure 1.8 représente sous forme d'histogramme la répartition des températures opératives sur la zone du rez-de-jardin pendant la période estivale. On observe une importante hausse de température dans le cas d'un albédo de 0,4 en comparaison à l'albédo de 0,2.



**Figure 1.8** – Diagramme de répartition des températures intérieures dans la maison INCAS-DM sur la période estivale en fonction des différents cas [18]

Le fort impact observé dépend de l'efficacité énergétique du bâtiment étudié ce qui montre l'importance de considérer avec attention le paramètre albédo. Une autre source d'incertitude est la modélisation des casquettes et on va étudier les conséquences de leur sous-évaluation sur les simulations .

### 1.2.3 Impact d'une hypothèse de simulation sur les masques proches

#### 1.2.3.1 Description des débords de fenêtre

L'hypothèse de simulation mise en cause est de considérer ou non ou partiellement les masques proches dans une simulation. Cette incertitude intervient lors de l'élaboration du modèle géométrique du bâtiment. L'utilisateur tente de représenter au mieux la réalité géométrique. Des hypothèses (approximations) de modélisation géométrique doivent être faites et elles dépendent fortement de l'utilisateur. La prise en compte de cette hypothèse influence l'impact des apports solaires (effet d'ombrage) d'où une augmentation prévisible de la surchauffe du bâti en été (inconfort thermique) si on ne modélise pas les masques proches. On peut s'attendre également à une sous-estimation des besoins de chauffage en hiver car on diminue les apports solaires "gratuits".

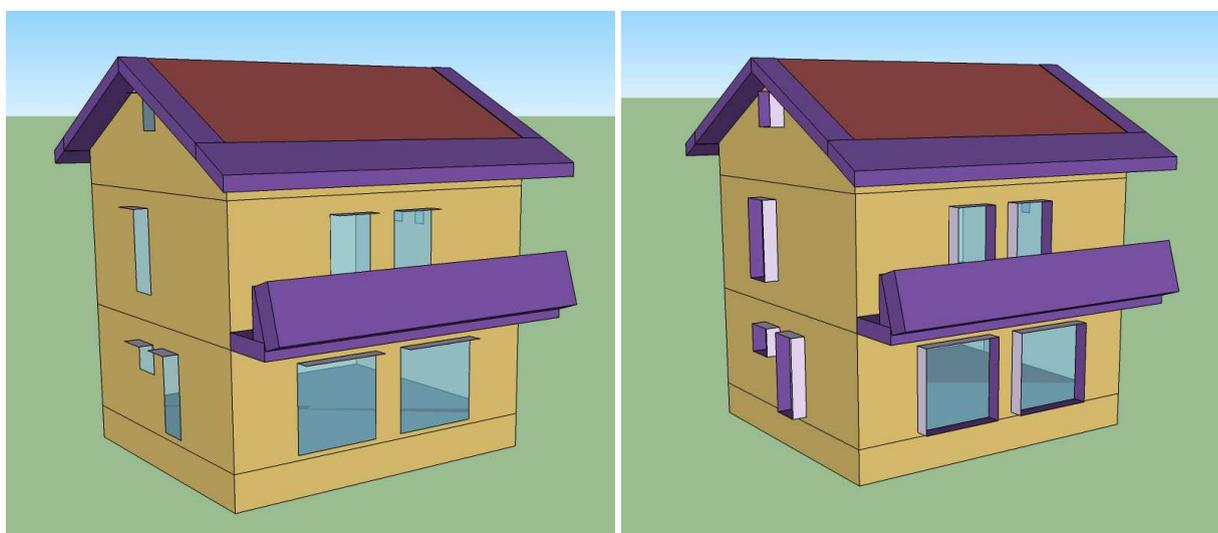
#### 1.2.3.2 Impact de l'hypothèse de simulation sur la maison INCAS-DM

L'épaisseur des murs de la maison étudiée est de 50 cm et les fenêtres sont placées au niveau de l'isolant et donc "décalées" de 35 cm par rapport au bord de façade (extérieure). La simulation ne considère que l'aspect thermique de l'épaisseur d'un mur et si on veut tenir compte des débords de fenêtre il faut les rajouter sur le modèle géométrique. Il se pose alors la question de prendre en compte ou non ces "bords de fenêtres" virtuels.

Deux simulations ont été effectuées avec EnergyPlus pour mesurer l'impact de ce "détail" géométrique avec un albédo de 0,2. Les captures écran des deux modèles géométriques testés sont présentées en figure 1.9 avec d'avantage de détails figure 1.10. Une configuration avec uniquement une casquette et une autre où la totalité du cadre est représentée (casquette, bords latéraux et bas).

Si on considère 27°C la température intérieure opérative au-delà de laquelle on suppose le bâtiment en surchauffe, on obtient pour la simulation B une valeur de 40 jours de surchauffe contre 85 jours dans le cas A (simulation sans bords latéraux de fenêtre). De la même façon que pour l'albédo, la figure 1.11 représente la répartition des températures opératives au rez-de-jardin dans les deux configurations testées.

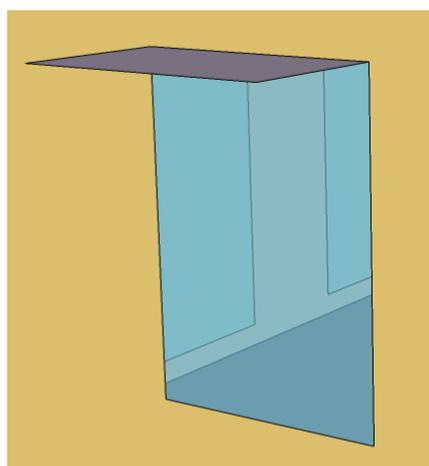
On observe de nouveau un décalage des températures extrêmes entre les deux cas. En négligeant l'hypothèse des débords complets de fenêtres on surestime l'inconfort d'été et donc les besoins de refroidissement. L'impact de la gestion des ombrages dans le modèle géométrique dépend des caractéristiques du bâtiment pris en compte. Si c'est un bâtiment à stratégie passive, l'impact est plus important que sur un bâtiment classique. Le



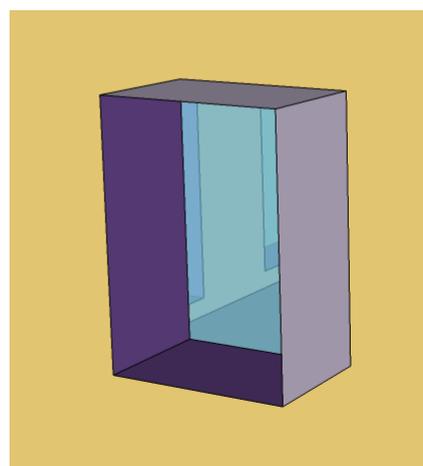
(a) Cas A

(b) Cas B

**Figure 1.9** – Les captures écran des deux cas simulés pour les fenêtres. A gauche le cas A : casquette seule et à droite le cas B : débords complets



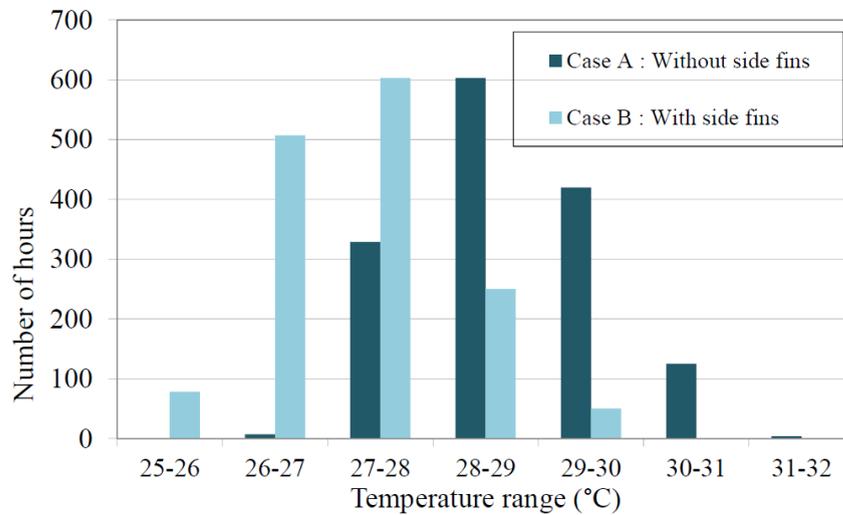
(a) Cas A



(b) Cas B

**Figure 1.10** – Type de masque proche sur les fenêtres pour le cas A : casquette seule et le cas B : débords complets

renforcement de fenêtre étudié ici est fréquent dans le cadre de réhabilitation thermique de bâtiment ancien isolé par l'extérieur.



**Figure 1.11** – Diagramme de répartition des températures intérieures dans la maison INCAS-DM sur la période estivale en fonction des différents cas, en bleu foncé le cas A casquette seule, en bleu clair le cas B débords complets [18]

### 1.3 Un besoin de transparence sur la gestion des sollicitations par le bâtiment basse consommation

Par les exemples ci-dessus, on observe l'influence importante d'une hypothèse de simulation et d'une entrée statique liées toutes deux à la gestion du rayonnement solaire par le bâtiment (réflexion du sol, masque proche). Ce type de bâtiment dit passif et optimisant les gains solaires en hiver et été est très sensible aux apports solaires, d'autant plus que le modèle thermique ne prend pas en compte de renouvellement d'air en plus des infiltrations. On souhaite obtenir plus de transparence sur sa capacité à gérer ces apports. Évaluer l'incertitude sur l'albédo ou sur la géométrie des ombrages est pertinent mais ce sont surtout les incertitudes liées aux conditions limites du modèle et notamment la météorologie sur laquelle il devient important de se questionner. En effet, nous avons vu que les bâtiments basse consommation en diminuant l'effet de la température extérieure, sont plus sensibles aux apports solaires. On se pose alors la question de l'impact de la variabilité météo sur l'incertitude de sortie ainsi que l'influence de chaque variable météo sur les besoins énergétiques.

La démarche visant l'amélioration de la fiabilité des résultats consiste à :

- Propager les incertitudes liées à la météo afin d'obtenir un intervalle de confiance sur la performance énergétique du bâtiment
- Identifier la ou les variables météorologiques responsables de la réponse du bâtiment.

Pour cela les techniques d'analyse de sensibilité sont utilisées et une stratégie de gestion est élaborée pour traiter la problématique des incertitudes et de l'influence des sollicitations météo sur la performance énergétique d'une maison basse consommation d'énergie. Avant de présenter cette stratégie, le formalisme mathématique des analyses de sensibilité est présenté ainsi que les bonnes pratiques associées.



# Chapitre 2

## Outils Mathématiques pour le gain en fiabilité : les analyses de sensibilité globale et d'incertitude

---

Dans ce chapitre, on présente le principe des analyses de sensibilité avant de détailler la démarche d'estimation de la sensibilité ainsi que des outils complémentaires permettant un calcul robuste et performant de la sensibilité adapté aux modèles complexes de grande dimension en simulation thermique dynamique. Pour terminer, on présente la difficulté de la gestion des entrées fonctionnelles en analyse de sensibilité.

---

### 2.1 Des outils puissants

#### 2.1.1 Rayons X du modèle conditionnés par un objectif

Les techniques d'analyse de sensibilité et d'analyse d'incertitude sont des atouts pour le gain en fiabilité et en confiance sur les résultats de la simulation thermique dynamique. L'analyse d'incertitude permet d'obtenir un intervalle de confiance sur la réponse du modèle en quantifiant l'amplitude de l'incertitude sur la sortie du modèle causée par les incertitudes en entrée, alors que l'analyse de sensibilité permet de déterminer les paramètres influents dans le code de calcul sondé par perturbation des entrées incertaines. Si les variations d'entrée correspondent à leurs incertitudes alors l'analyse de sensibilité et d'incertitude se rejoignent en une seule analyse. Ainsi deux types d'information sont extraites d'une analyse des performances énergétiques d'un bâtiment basse consommation par la simulation thermique dynamique : l'intervalle de confiance sur la performance ainsi

que l'identification des entrées auxquelles les réponses du modèle sont sensibles. Ces informations peuvent permettre d'identifier les paramètres sur lesquels il est pertinent de quantifier les incertitudes et de proposer des priorités de mise à jour du modèle.

Il faut vérifier la bonne évaluation de ces entrées influentes et des modèles attenants. De la même manière, si des entrées ne sont pas influentes, on pourra plus aisément fixer leur valeur et simplifier le modèle même si l'analyse de sensibilité en fixant certaine valeur peut induire un biais sur les indices de sensibilité. L'estimation de cette erreur est décrite par Sobol et al. [59].

Selon Lilburne et al. [36], l'analyse de sensibilité doit suivre les six étapes suivantes :

1. Définir l'objectif de l'étude
2. Sélectionner les entrées que l'on veut étudier
3. Associer un intervalle d'incertitude et une loi de probabilité
4. Appliquer une méthode de génération pour obtenir les échantillons de N simulations à partir des distributions des entrées
5. Réaliser les simulations pour chaque jeu d'échantillons afin d'obtenir les N valeurs de la sortie de simulation
6. Utiliser les résultats de l'étape 5 pour l'analyse d'incertitude et le calcul des indices de sensibilité pour quantifier l'effet de chaque entrée sur la sortie

L'analyse de sensibilité dépend de l'objectif de l'étude car il conditionne le choix de l'indice de sensibilité approprié et de la stratégie de l'échantillonnage. De plus l'interprétation des résultats dépend de la question posée en amont à laquelle l'analyse de sensibilité doit répondre. Plus la question est ciblée et plus l'interprétation aura des chances d'apporter une réponse précise rendant les conclusions d'autant plus efficaces. Dans le cadre des problématiques associées à la thermique du bâtiment et au vu de la multitude d'aspect pouvant être étudiés, il est primordial de bien délimiter les attentes liées à l'analyse de sensibilité. L'étude de l'impact de la forme ou des volumes sur la performance énergétique (Hopfe et al. [25]) n'inclut pas les mêmes entrées incertaines qu'une étude sur un bâtiment préexistant où la géométrie est fixe. De même, les choix des plages de variation ainsi que de la loi de probabilité associée doivent être adaptés à l'objectif. Le choix de l'amplitude des variations et de la zone du modèle étudié impactent de manière significative les résultats obtenus. Les deux exemples ci-dessous illustrent ces notions.

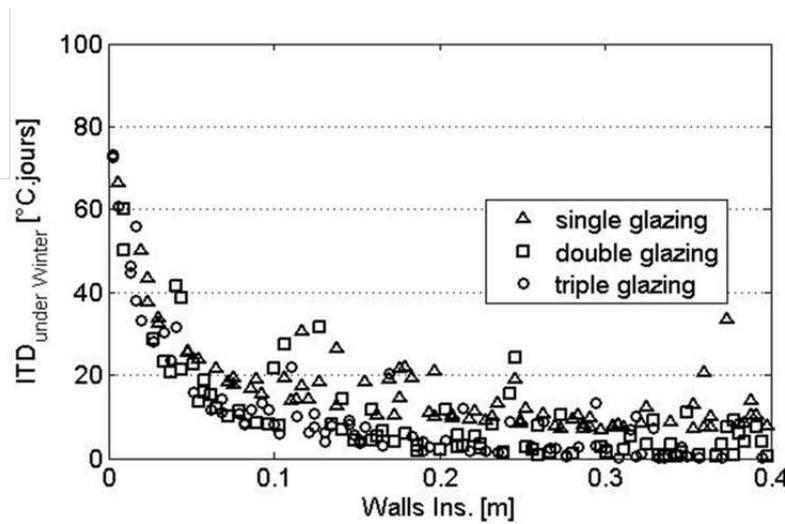
### 2.1.2 Plage de variation et zone d'exploration du modèle : cas de la thermique du bâtiment

Pour illustrer ces aspects deux problématiques d'analyse de sensibilité associées au même cas test (section 1.2.1) sont confrontées. L'une traite de la problématique de la validation de modèle (Spitz et al. [62]) et l'autre de l'optimisation pour la réhabilitation (Rabouille et al. [52]). L'objectif de ces deux analyses de sensibilité diffère.

Pour la validation de modèle, les informations que l'on souhaite obtenir de l'analyse de sensibilité sont une plage d'incertitude de simulation ainsi que des paramètres influents. Le but de l'étude étant de valider un modèle, une incertitude de simulation doit permettre de le comparer aux incertitudes de mesure. Par ailleurs les paramètres influents sont déterminés pour vérifier la bonne estimation de leur incertitude et la possibilité de calibrer le modèle dans la plage cohérente du paramètre.

En ce qui concerne la problématique liée à la réhabilitation, les paramètres influents sont déterminés pour élaborer un plan de rénovation énergétique par optimisation des performances. La différence entre ces deux études se situe principalement dans la plage de variation imposée aux entrées. Dans le cas de la validation, les incertitudes doivent être cohérentes avec le site expérimental. Nous avons mentionné en chapitre 1, la modification de la valeur d'albédo entre les plans de conception et la mise en œuvre réelle (gravier blanc à la place de bitume). Nous avons également montré que le bâtiment de type passif était significativement sensible à la modification de ce paramètre.

La caractérisation des incertitudes en thermique du bâtiment est une problématique délicate à laquelle peu de travaux ont été consacrés ([9], [37]). Il n'est pas forcément évident de différencier les incertitudes dites naturelles du paramètre de celles liées à la méconnaissance. Cette méconnaissance peut être due aux phénomènes physiques en interaction (humidité, vieillissement) ou bien à la mise en œuvre (modification de plan, mauvaise pose d'isolant). Pour les incertitudes naturelles, nous avons vu en chapitre 1, que la loi normale est la plus adaptée. Lorsque les incertitudes sont régies par un phénomène sous-jacent et que la loi de distribution n'est pas identifiable, il est préférable de s'orienter vers une loi uniforme. La loi uniforme va représenter une équiprobabilité d'obtenir une valeur entre deux bornes. Dans le cas de la problématique de la réhabilitation cette loi est tout indiquée. En effet, les plages de variation des entrées correspondent aux opérations de rénovation envisageables et si aucune contrainte financière en amont n'est prise en compte dans l'étude thermique, le test du modèle avec une isolation de 10 cm que de 30 cm à la même probabilité. Ces plages sont beaucoup plus larges qu'une simple incertitude et permettent d'explorer le modèle sur une zone étendue de leur comportement.



**Figure 2.1** – Evolution de la variable de sortie inconfort d’hiver en fonction de la variabilité générée de l’entrée "épaisseur d’isolant" [52]

La figure 2.1 montre un exemple de l’évolution d’un des indicateurs, l’inconfort thermique en hiver, en fonction d’un des paramètres influents de l’étude, l’épaisseur d’isolant (Scatterplot). Nous observons sur la figure une large gamme de variation. Ce qui est intéressant c’est qu’en fonction de l’endroit où le modèle est évalué, l’influence du paramètre ne sera pas la même. Si on opère une variation entre 0 et 10 cm l’influence sera plus importante qu’entre 30 et 40 cm. Ce type d’analyse est fait pour permettre la réduction de modèle en établissant les liens directs entre indicateurs de performance et paramètres influents.

Le but de ces deux exemples d’analyse de sensibilité relativement fréquents en thermique du bâtiment est d’illustrer la problématique associée à la définition des plages d’incertitudes ainsi que l’importance d’interpréter les résultats des indices de sensibilité en prenant bien en compte le contexte de la simulation, c’est-à-dire l’exploration du modèle. Les paramètres influents sur un modèle de bâtiment énergivore ne sont pas les mêmes que sur un modèle de bâti basse consommation comme le montre Spitz et al. [61].

Cette réflexion intervient dans un contexte où les méthodes d’analyse de sensibilité deviennent de plus en plus accessibles<sup>1</sup>, et des modèles complémentaires<sup>2</sup> associés aux logiciels de simulation thermique dynamique permettent l’automatisation d’analyse de sensibilité. Il faut s’assurer d’avoir un objectif en amont de la démarche et de bien définir

<sup>1</sup>Développement de toolbox Matlab et R, outils comme Simlab et OpenTurns

<sup>2</sup>jEPlus pour EnergyPlus

le contexte de l'étude. Ce contexte va également conditionner le type de méthode que l'on va employer.

### 2.1.3 Choix de la méthode

Le choix de la technique d'analyse de sensibilité selon Saltelli et al. [58] dépend de plusieurs points :

- le temps d'une simulation
- le nombre d'entrées incertaines
- les hypothèses du modèle (linéaire, monotone, etc..)
- la présence d'interactions ou non dans le modèle et donc la nécessité de les prendre en compte
- le cadre (contexte de l'analyse) et le public visé (interprétation plus ou moins aisée en fonction des indices de sensibilité)

La simulation thermique dynamique étant non linéaire, on élimine les analyses de sensibilité locale qui quantifient l'impact sur la sortie du modèle lorsque des petites perturbations sont appliquées aux entrées incertaines. Cette quantification dépend des dérivées partielles qui sont calculées à un point bien précis de l'espace des paramètres. Le choix du point peut influencer le résultat comme nous avons pu le constater avec l'exemple de la réhabilitation, surtout si cet espace est important (plage de variation significatives). Ainsi, l'information extraite ne vaut qu'au niveau du point de l'espace des paramètres où la sensibilité a été estimée et ne permet pas de conclure sur le reste de l'espace des paramètres ce qui n'est pas un problème pour les modèles linéaires où une extrapolation peut permettre de calculer rapidement les indices à différents points de l'espace mais dans le cas de modèle non linéaire l'information est uniquement ponctuelle donc limitée. Les analyses de type "un facteur à la fois"(one-at-the-time : OAT) font aussi partie de ces méthodes où une part importante de l'espace des paramètres n'est pas explorée. Même si l'analyse est faite sur des plages d'incertitudes conséquentes, la non simultanité des variations sur les entrées entraîne la non prise en compte des interactions, ce qui peut représenter plus de 99% d'espace des paramètres inexploré (Saltelli et al. [56]).

Nous allons nous concentrer sur les différents type d'analyse de sensibilité globale qui permettent un retour détaillé des relations entre les paramètres du modèle, interactions comprises. Selon Sudret [63], il existe deux types de méthode :

- celles basées sur la régression : pour ce type d'approche, une régression linéaire lie les entrées et les sorties obtenues par les points des simulations Monte-Carlo

(échantillonnage aléatoire sur l'espace des paramètres) et le coefficient est obtenu par la technique des moindres carrés. Dans le cas général de modèle non-linéaire et non-monotone, cette approche ne donne pas de résultat satisfaisant [58].

- celles basées sur la variance : ce genre de technique peut calculer les indices de sensibilité indépendamment des hypothèses du modèle. Dans les méthodes basées sur la variance, la variance  $V$  de la sortie du modèle est décomposée en variance partielle qui représente la part de  $V$  qui est expliquée par les entrées du modèle (que l'on considère une entrée seule, ou un groupe d'entrées) et par conséquent l'influence des entrées elles-mêmes sur la sortie. On distingue deux approches communément utilisées : la méthode FAST (pour Fourier amplitude sensitivity test), à chaque entrée est associée une fréquence de variation, l'analyse fréquentielle de la sortie va permettre de déterminer les paramètres influents. Une autre approche consiste à calculer les indices de Sobol mais peut être coûteuse en évaluation de modèle.

Il existe une quantité importante de méthodes qui ont été développées pour permettre l'analyse de sensibilité et le choix dépend du cas d'application, les travaux de Saltelli et al. [58] et [55] regroupent les références et les notions permettant de guider au mieux les utilisateurs.

Dans le cadre de ce travail, la méthode sélectionnée est basée sur la variance, il s'agit des indices de sensibilité de Sobol [60]. Ces indices sont indiqués pour les modèles complexes tels que ceux impliqués dans la simulation thermique dynamique. Ils permettent la gestion d'entrées de nature différente (catégorique, fonctionnelle, scalaire). Leur interprétation est intuitive, les indices sont normalisés et compris entre 0 et 1. Par ailleurs, il est possible d'estimer des indices de sensibilité sur des groupes de paramètres, ce qui est un élément clé dans la stratégie de gestion des entrées fonctionnelles développées au chapitre 3. L'estimation des indices peut être coûteuse en évaluation de modèles de grande dimension, mais cet inconvénient peut toutefois être atténué grâce à une procédure d'échantillonnage présentée ci-après en section 2.2.3.1.

La partie suivante décrit la procédure et le formalisme associé à la démarche de calcul des indices de sensibilité de Sobol.

## 2.2 Des indices de sensibilité adaptés à la problématiques du bâtiment

### 2.2.1 Des indices basés sur la variance

Soit  $X = \{X_1, \dots, X_i, \dots, X_p\}$  l'ensemble des entrées incertaines du modèle et  $Y$  la sortie, que l'on suppose scalaire. Pour obtenir l'influence de chaque  $X_i$  sur la sortie  $Y$ , on utilise les méthodes probabilistes de type Monte-Carlo, afin d'échantillonner les entrées en les faisant varier chacune indépendamment sur leur loi de probabilité donc sur leur plage d'incertitude. On obtient ainsi des matrices de dimension  $N \times p$  avec  $N$  le nombre de tir et  $p$  le nombre d'entrées. Pour pouvoir quantifier l'effet de  $X_i$  sur  $Y$ , on évalue la part de la variance  $Y$  induite par  $X_i$ . Cette quantité est définie par Sobol [60] comme étant l'indice de sensibilité de premier ordre :

$$S_i = \frac{V(E[Y|X_i])}{V(Y)} \quad (2.1)$$

avec  $V(\cdot)$  la variance et  $E(\cdot)$  l'espérance mathématique. Cet indice a été introduit par décomposition de la fonction  $f$  du modèle en somme de fonctions de dimensions croissantes :

$$Y = f(X_1, \dots, X_p) = f_0 + \sum_{i=1}^p f_i(X_i) + \sum_{1 \leq i < j < p} f_{ij}(X_i, X_j) + \dots + f_{1\dots p}(X_1, \dots, X_p) \quad (2.2)$$

où

$$\begin{aligned} f_0 &= E[Y], \\ f_i(X_i) &= E[Y|X_i] - E[Y], \\ f_{ij}(X_i, X_j) &= E[Y|X_i, X_j] - E[Y|X_i] - E[Y|X_j] + E[Y] \end{aligned}$$

et ainsi de suite pour les ordres supérieurs.

Une telle décomposition est unique, si et seulement si, les variables aléatoires sont indépendantes et que les fonctions de la décomposition (2.2) sont orthogonales. La variance se décompose alors ([60]) :

$$V = \sum_i V_i + \sum_{i < j} V_{ij} + \sum_{i < j < m} V_{ijm} + \dots + V_{12\dots k} \quad (2.3)$$

où

$$\begin{aligned} V_i &= V(E[Y|X_i]), \\ V_{ij} &= V(E[Y|X_i, X_j]) - V_i - V_j, \\ V_{ijk} &= V(E[Y|X_i, X_j, X_k]) - V_{ij} - V_{ik} - V_{jk} - V_i - V_j - V_k, \\ &\vdots \\ V_{1\dots p} &= V - \sum_{i=1}^p V_i - \sum_{1 \leq i < j \leq p} V_{ij} - \dots - \sum_{1 \leq i < \dots < i_{p-1} \leq p} V_{i_1 \dots i_{p-1}} \end{aligned}$$

Si on divise chaque membre de l'équation 2.3 par  $V(Y)$ , on obtient :

$$\sum_i S_i + \sum_{i < j} S_{ij} + \sum_{i < j < m} S_{ijm} + \dots + S_{12\dots k} = 1 \quad (2.4)$$

Avec

$$S_{ij} = \frac{V_{ij}}{V(Y)} \quad (2.5)$$

l'indice de sensibilité d'ordre deux, c'est-à-dire la sensibilité de  $Y$  aux variables  $X_i$  et  $X_j$ , ce qui correspond à leur interaction. Cet effet n'est pas pris en compte dans l'effet du première ordre qui quantifie l'effet des variables seules.

De même, l'effet d'ordre trois :

$$S_{ijk} = \frac{V_{ijk}}{V(Y)} \quad (2.6)$$

exprime l'effet d'interaction des variables  $X_i$ ,  $X_j$  et  $X_k$  sur  $Y$ . Ces effets ne sont pas pris en compte dans les indices de premier et second ordre. Et de manière similaire jusqu'à l'ordre  $p$ . L'indice de sensibilité totale  $S_T$  introduit par Homma et al. [23], regroupe l'ensemble des effets relatifs à la variable  $X_i$  :

$$S_{T_i} = \sum_{k \# i} S_k \quad (2.7)$$

où  $\#i$  représente tous les ensembles d'indices contenant l'indice  $i$ .

Étant normalisés, les indices sont compris entre 0 et 1, plus l'indice est grand et plus l'effet de la variable sur la sortie est important. La somme des indices de sensibilité est égale à 1 (2.4). Ainsi si on vérifie que  $\sum_i S_i \sim 1$ , alors les effets d'ordre supérieur à 1 peuvent être négligés puisque peu influents. La procédure est donc d'obtenir dans un premier temps l'effet seul de chaque variable par le calcul de  $S_i$ , puis, en fonction de la somme des  $S_i$  la présence ou non d'interactions est facilement déterminée et conditionne le calcul ou non des effets totaux.

Les indices de sensibilité de Sobol sont intuitifs et particulièrement recommandés dans le cas de modèles complexes de grande dimension de système incertain comme la thermique du bâtiment. La partie suivante détaille une des procédures pour les estimer.

## 2.2.2 Démarche et bonnes pratiques pour l'obtention des indices

Soit un code de calcul représenté par  $f(\cdot)$ ,  $Y$  la sortie scalaire et  $\mathbf{X} = (X_1, \dots, X_p)$  les entrées du modèle, on a :  $Y = f(X_1, \dots, X_p)$ . Soit  $N$  le nombre de tirs aléatoires pour l'échantillonnage des entrées, on obtient une variable  $Y$  composée des  $N$  réponses du modèle. Chaque simulation donne une valeur de  $Y$ . Une simulation correspond à une

évaluation du modèle à partir d'un jeu de  $p$  entrées généré aléatoirement en fonction de la distribution de chaque entrée 2.8.

$$f\left(\begin{bmatrix} X_{1,1} & \dots & X_{1,i} & \dots & X_{1,p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{j,1} & \dots & X_{j,i} & \dots & X_{j,p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{N,1} & \dots & X_{N,i} & \dots & X_{N,p} \end{bmatrix}\right) = \begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_N \end{bmatrix} \quad (2.8)$$

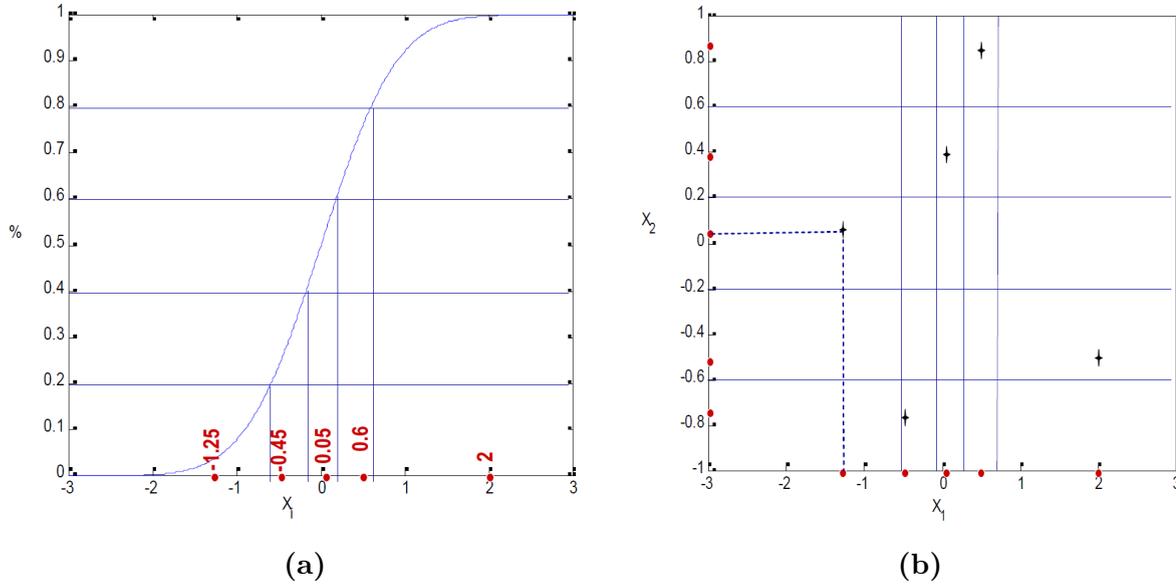
Le bon calcul des indices de sensibilité dépend de l'estimateur mais aussi de l'échantillonnage, les deux prochaines sous-sections traitent de ces aspects.

### 2.2.2.1 Échantillonnage par hypercube latin

L'échantillonnage des variations d'entrées conditionne les résultats d'analyse de sensibilité. La taille de l'échantillon et la manière d'échantillonner jouent sur la précision et la rapidité de l'estimation des indices. Le balayage de l'espace des paramètres d'entrée doit être homogène pour s'assurer de la bonne représentativité des incertitudes d'entrées. La méthode la plus communément utilisée est le tir aléatoire classique. Cette méthode peut être optimisée afin de converger plus rapidement vers l'estimation des indices de sensibilité. Pour cela l'espace des paramètres est divisé en sous-ensembles pour permettre un échantillonnage plus représentatif en moins de tir. Une de ces méthodes est le Latin Hypercube Sampling (LHS) ou échantillonnage par le latin hypercube qui permet d'obtenir à moindre coût de simulation des résultats d'indice de sensibilité plus stables (Helton et al. [22]). La méthode consiste dans un premier temps à diviser l'intervalle de variations de chaque entrée en  $n_p$  sous-intervalles équiprobables. Puis une valeur est tirée aléatoirement dans chaque sous-intervalle. Cette procédure permet, même avec un nombre limité de tir, une représentation de l'espace des paramètres.

La figure 2.2 illustre cette procédure, avec  $X_1$  et  $X_2$  deux paramètres dont l'un a une loi de probabilité normale centrée réduite, représentée par la fonction de distribution en figure 2.2a. Les 5 sous-intervalles équiprobables  $y$  sont délimités. Sur la figure 2.2b, les 5 couplets générés sont représentées.

D'autres méthodes déterministes d'échantillonnage de type quasi Monte-Carlo permettent également une convergence plus rapide qu'avec LHS, voir Saltelli et al. [58] pour plus de détails. Le principe de ces méthodes tient compte des autres points générés. Cependant nous nous limiterons dans ce travail à l'utilisation de Latin hypercube pour sa facilité d'implémentation et pour la possibilité de générer des échantillons corrélés [27]. Cette particularité est exploitée dans le cadre de la gestion des entrées temporelles et est



**Figure 2.2** – Illustration du principe de l'échantillonnage hypercube latin. En a) le découpage des 5 sous-intervalles selon la fonction de distribution d'une des entrées, en b) représentation des 5 couplets générés [38]

discutée en chapitre 3.

### 2.2.2.2 Une estimation précise

Il s'agit désormais de déterminer de manière la plus précise possible les indices de sensibilité. Pour cela on utilise des estimateurs. Le résultat doit être stable, indépendant du modèle dans l'idéal, être précis (faible variance) et proche de l'exactitude (faible biais). La méthode présentée ci-dessous est l'une des plus robuste et précise pour le calcul des indices de sensibilité de Sobol et est développée par Saltelli et al. [57].

On génère deux matrices d'échantillons  $N \times p$  de manière indépendante. Appelons  $\mathbf{A}$  et  $\mathbf{B}$  ces deux matrices, avec respectivement  $a_{ji}$  et  $b_{ji}$  les éléments les composant. L'indice  $i$  est associé au nombre  $p$  d'entrée et l'indice  $j$  est associé au nombre de simulations  $N$  (cf matrice 2.8). De ces deux matrices  $\mathbf{A}$  et  $\mathbf{B}$ , on construit  $\mathbf{A}_{\mathbf{B}}^{(i)}$  telle que toutes les colonnes proviennent de  $\mathbf{A}$  sauf la  $i$ -ième colonne qui est prise dans  $\mathbf{B}$ . De même la matrice  $\mathbf{B}_{\mathbf{A}}^{(i)}$  est construite où toutes les colonnes proviennent de  $\mathbf{B}$  sauf la  $i$ -ième colonne qui est prise dans  $\mathbf{A}$ . On a :

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,i} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,i} & \dots & a_{2,p} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ a_{N,1} & a_{N,2} & \dots & a_{N,i} & \dots & a_{N,p} \end{bmatrix} \quad (2.9)$$

$$\mathbf{B} = \begin{bmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,i} & \dots & b_{1,p} \\ b_{2,1} & b_{2,2} & \dots & b_{2,i} & \dots & b_{2,p} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ b_{N,1} & b_{N,2} & \dots & b_{N,i} & \dots & b_{N,p} \end{bmatrix} \quad (2.10)$$

Les matrices hybrides :

$$\mathbf{B}_{\mathbf{A}}^{(i)} = \begin{bmatrix} b_{1,1} & b_{1,2} & \dots & a_{1,i} & \dots & b_{1,p} \\ b_{2,1} & b_{2,2} & \dots & a_{2,i} & \dots & b_{2,p} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ b_{N,1} & b_{N,2} & \dots & a_{N,i} & \dots & b_{N,p} \end{bmatrix} \quad (2.11)$$

et

$$\mathbf{A}_{\mathbf{B}}^{(i)} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & b_{1,i} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & b_{2,i} & \dots & a_{2,p} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ a_{N,1} & a_{N,2} & \dots & b_{N,i} & \dots & a_{N,p} \end{bmatrix} \quad (2.12)$$

Seule la valeur de  $X_i$  reste inchangée entre  $\mathbf{B}$  et  $\mathbf{B}_{\mathbf{A}}^{(i)}$ . A partir du couple de matrice  $\mathbf{A}, \mathbf{B}_{\mathbf{A}}^{(i)}$  ou bien  $\mathbf{B}, \mathbf{A}_{\mathbf{B}}^{(i)}$  les  $S_i$  peuvent être calculés, selon [60] :

$$V(E[Y|X_i]) = \frac{1}{N} \sum_{j=1}^N f(\mathbf{A})_j f(\mathbf{B}_{\mathbf{A}}^{(i)})_j - f_0^2 \quad (2.13)$$

Une amélioration de cet estimateur est proposé par Saltelli et al.[57] :

$$V(E[Y|X_i]) = \frac{1}{N} \sum_{j=1}^N f(\mathbf{B})_j (f(\mathbf{A}_{\mathbf{B}}^{(i)})_j - f(\mathbf{A})_j) \quad (2.14)$$

On obtient bien l'espérance de  $Y$  connaissant  $X_i$  sachant que seule l'entrée  $X_i$  est identique entre les matrices. La variance de cette quantité divisée par la variance totale de la sortie nous donne  $S_i$ .

Le nombre d'évaluations du modèle  $N_T$  est de  $N \times (p + 2)$  à  $N \times p$  pour calculer les  $S_i$  avec  $N$  devant être au moins égal à 500 (Saltelli et al. [57]). Nous verrons que dans le cadre de modèles de thermique du bâtiment le nombre d'entrées  $p$  est important et le nombre  $N_T$  d'évaluations du modèle peut vite exploser. C'est la raison pour laquelle nous utiliserons la méthode développée par Mara et al [39] afin de rendre le nombre  $N_T$  indépendant de  $p$ . Cette méthode est présentée ci-après.

### 2.2.3 Outils complémentaires : intervalle de confiance et réduction du coût de simulation pour les $S_i$

#### 2.2.3.1 Réduction drastique du nombre de simulation par permutation de matrice

Cette méthode [39] permet de rendre le nombre  $N_T$  d'évaluation totale du modèle indépendant du nombre  $p$  d'entrées. Pour cela, on génère une matrice  $\mathbf{X}^{(0)}$  (la matrice de référence) de dimension  $N \times p$ , puis on construit une seconde matrice à partir de  $\mathbf{X}^{(0)}$  en permutant aléatoirement chaque colonne  $i$  de  $\mathbf{X}^{(0)}$ , on appelle cette seconde matrice  $\mathbf{X}^{(1)}$ .

Ci-dessous, une illustration du principe, avec des matrices d'échantillon de  $p = 3$  entrées et  $N = 5$  jeux de simulations. On génère aléatoirement sur une loi uniforme comprise entre 0 et 1 soit  $\mathcal{U}[0, 1]$  les cinq jeux de simulation pour les trois variables. Soit, par exemple,

$$\mathbf{X}^{(0)} = \begin{bmatrix} 0,75 & 0,53 & 0,47 \\ 0,38 & 0,78 & 0,01 \\ 0,57 & 0,93 & 0,34 \\ 0,08 & 0,13 & 0,16 \\ 0,05 & 0,57 & 0,79 \end{bmatrix} \quad (2.15)$$

Dont la matrice des rangs associée est :

$$\mathbf{R}_0 = \begin{bmatrix} 5 & 2 & 4 \\ 3 & 4 & 1 \\ 4 & 5 & 3 \\ 2 & 1 & 2 \\ 1 & 3 & 5 \end{bmatrix} \quad (2.16)$$

On génère à partir de  $\mathbf{X}^{(0)}$  un second échantillon indépendant de matrice des rangs,

$$\mathbf{R}_1 = \begin{bmatrix} 4 & 2 & 4 \\ 1 & 3 & 1 \\ 5 & 4 & 5 \\ 3 & 5 & 3 \\ 2 & 1 & 2 \end{bmatrix} \quad (2.17)$$

Notons  $\mathbf{X}^{(1)} = \mathbf{R}_1(\mathbf{X}^{(0)})$  une telle transformation. On obtient,

$$\mathbf{X}^{(1)} = \begin{bmatrix} 0,57 & 0,53 & 0,47 \\ 0,05 & 0,57 & 0,01 \\ 0,75 & 0,78 & 0,79 \\ 0,38 & 0,93 & 0,34 \\ 0,08 & 0,13 & 0,16 \end{bmatrix} \quad (2.18)$$

On évalue le modèle pour les deux jeux d'échantillons  $\mathbf{X}^{(0)}$  et  $\mathbf{X}^{(1)}$ , tel que

$$\mathbf{Y}^{(0)} = f(\mathbf{X}_1^{(0)}, \mathbf{X}_2^{(0)}, \dots, \mathbf{X}_p^{(0)}); \mathbf{Y}^{(1)} = f(\mathbf{X}_1^{(1)}, \mathbf{X}_2^{(1)}, \dots, \mathbf{X}_i^{(1)}, \dots, \mathbf{X}_p^{(1)})$$

Pour permettre le calcul des  $S_i$  on doit obtenir deux jeux de sorties où seul  $\mathbf{X}_i$  reste inchangé comme vu avec l'estimateur présenté en section 2.2.2.2. Avec  $\mathbf{Y}^{(0)}$  la variable de sortie de référence on appelle  $\mathbf{Y}_i^{(2)}$  la variable de sortie liée à un échantillon d'entrée où toutes les valeurs proviennent de  $\mathbf{Y}^{(1)}$  sauf  $\mathbf{X}_i$  qui provient de  $\mathbf{Y}^{(0)}$ , tel que :

$$\mathbf{Y}^{(0)} = f(\mathbf{X}_1^{(0)}, \mathbf{X}_2^{(0)}, \dots, \mathbf{X}_i^{(0)}, \dots, \mathbf{X}_p^{(0)}); \mathbf{Y}^{(2)} = f(\mathbf{X}_1^{(1)}, \mathbf{X}_2^{(1)}, \dots, \mathbf{X}_i^{(0)}, \dots, \mathbf{X}_p^{(1)})$$

Notons  $\mathbf{RP}_i$  le vecteur correspondant à la  $i$ -ème colonne de la matrice des rangs  $\mathbf{R}_0$ . Alors, la transformation recherchée est  $\mathbf{Y}_i^{(2)} = \mathbf{RP}_i(\mathbf{Y}^{(0)})$ . Cette permutation équivaut à établir la variable de sortie d'une matrice d'entrées où seule la valeur de  $X_i$  est identique avec  $Y^{(0)}$ , mais ne nécessite aucune simulation supplémentaire. Ainsi les indices de sensibilité  $S_i$  se calculent avec seulement  $2N$  simulations quelque soit le nombre  $p$  d'entrées.

L'estimateur :

$$V(E[Y|X_i]) = \frac{1}{N} \sum_{j=1}^N f(\mathbf{B})_j (f(\mathbf{A}_\mathbf{B}^{(i)})_j - f(\mathbf{A})_j) \quad (2.19)$$

devient

$$V(E[Y|X_i]) = \frac{1}{N} \sum_{j=1}^N Y_j^{(0)} (RP_i(Y^{(1)})_j - Y_j^{(1)}) \quad (2.20)$$

### 2.2.3.2 Quantification de la variabilité sur les indices de sensibilité par bootstrap

Pour associer à l'estimation des indices de sensibilité un intervalle de confiance, et cela sans évaluation du modèle supplémentaire, on utilise les méthodes de bootstrap [16].

Le bootstrap, ou technique de ré-échantillonnage avec remise, est une manière de refaire une analyse sans relancer de simulations. Il s'agit, à partir d'un échantillon statistique "source" représentatif, de recréer des échantillons de la même taille, pour cela on tire au hasard avec remise dans l'échantillon source parmi toutes les répliques possibles, sans évidemment toutes les construire (16796 répliques possibles pour une taille d'échantillons de 10). Puis, la variabilité de l'estimateur est évaluée par la variabilité observée sur l'ensemble des échantillons artificiels obtenus par ré-échantillonnage. Le bootstrap s'applique quelque soit la complexité de l'estimateur.

Principe du ré-échantillonnage bootstrap :

1. on crée des échantillons "artificiels" à partir de l'échantillon réel : appelés "répliques bootstrap"

2. on calcule ensuite l'estimateur (ici l'indice de sensibilité) sur chacun des échantillons artificiels (répliques)
3. on construit la loi de distribution empirique de la statistique (la distribution bootstrap)
4. on utilise cette distribution de la statistique pour estimer les propriétés de l'échantillon

Voici une illustration du principe avec un échantillon  $x$  de taille  $n = 5$  et l'estimateur moyenne  $\bar{x}$ . On génère  $k$  répliques bootstrap et  $k$  valeurs de moyenne estimées.

Soit :

$$x = \{19, 17, 26, 21, 22\} \quad \bar{x} = 21$$

Les  $k$  répliques bootstrap et donc l'estimation de la moyenne peuvent être les suivantes :

$$\begin{aligned}
 x^1 &= \{22, 17, 22, 19, 21\} & \bar{x}^1 &= 20, 2 \\
 x^2 &= \{22, 19, 22, 21, 19\} & \bar{x}^2 &= 20, 6 \\
 x^3 &= \{17, 22, 26, 26, 17\} & \bar{x}^3 &= 21, 6 \\
 & \vdots & & \vdots \\
 x^k &= \{22, 26, 17, 21, 22\} & \bar{x}^k &= 21, 6
 \end{aligned} \tag{2.21}$$

Le bootstrap permet d'évaluer la variabilité de l'estimateur et de l'échantillonnage mais aussi son biais. Le biais est l'écart entre la valeur calculée avec l'échantillon source et la moyenne de la distribution de la statistique évaluée par les répliques.

Soit l'estimateur de la statistique à partir de l'échantillon mère :

$$\hat{\theta} = f(x_1, x_2, \dots, x_n) \tag{2.22}$$

On suppose un bootstrap pour  $k$  répliques de l'échantillon, l'estimateur sur la  $b$ -ème réplique est telle que :

$$\hat{\theta}_{(b)} = f(x_1^*, x_2^*, \dots, x_n^*) \tag{2.23}$$

On a l'estimation bootstrap :

$$\hat{\theta}_{Boot} = \frac{1}{k} \sum_{b=1}^k \hat{\theta}_{(b)} \tag{2.24}$$

La variance de l'estimation bootstrap

$$Var_{Boot}(\hat{\theta}) = \frac{1}{k-1} \sum_{b=1}^k (\hat{\theta}_{(b)} - \hat{\theta}_{Boot})^2 \tag{2.25}$$

L'écart-type

$$\sigma_{Boot}(\hat{\theta}) = \sqrt{Var_{Boot}(\hat{\theta})} \quad (2.26)$$

Le bootstrap classique ne peut fonctionner en présence de corrélation dans l'échantillon, la corrélation serait brisée, voir Kreiss et al. [33] pour un état de l'art des adaptations bootstrap pour les échantillons corrélés.

Dans le cadre de ce travail de thèse le bootstrap va nous permettre d'estimer la variabilité des indices de sensibilité, en ré-échantillonnant les variables de sortie  $\mathbf{Y}^{(0)}$  et  $\mathbf{Y}^{(1)}$ . De par leur lien établi en section 2.2.3.1, une attention particulière doit être accordée dans le ré-échantillonnage de ces variables pour le calcul de  $S_i$ .

## **2.3 Gestion des entrées fonctionnelles en analyse de sensibilité**

Nous avons présenté une méthode d'estimation robuste des indices de sensibilité pour des modèles complexes comprenant des interactions. En complément, des outils ainsi que des astuces permettent d'améliorer les conditions d'obtention des résultats et l'intervalle de confiance. La discussion suivante porte sur la gestion en analyse de sensibilité globale des entrées de type temporel dans le cadre de modèles complexes à haute dimension.

### **2.3.1 Une problématique actuelle en thermique du bâtiment**

L'analyse de sensibilité avec entrées statiques indépendantes est une pratique bien délimitée et validée, que ce soit dans les méthodes ou même dans l'application en thermique du bâtiment, voir Tian et al. [65] pour un état de l'art à ce sujet. La prise en compte d'entrées corrélées, par contre, amène une difficulté supplémentaire que les méthodes standard ne peuvent résoudre. Des développements récents (Chastaing et al. [7], Kucherenko et al. [34]) permettent leurs gestions en adaptant les indices de Sobol mais l'interprétation devient délicate et l'implémentation de ces méthodes n'est pas évidente. La gestion des entrées corrélées est une problématique actuelle dans le développement des méthodes d'analyse de sensibilité.

Dans le cadre de la thermique du bâtiment, il est fréquent que les entrées soient corrélées, ces corrélations sont synonymes de liens entre les entrées, et ne pas les prendre en compte peut fausser les conclusions émergentes de l'analyse. On peut citer l'exemple de la corrélation entre l'âge des bâtiments et leur surface. En fonction des époques constructives, différentes typologies émergent. Ainsi les volumes d'un bâtiment sont fortement dépendants de leur âge. De même, une surface au sol peut être associée au nombre d'étage.

Ainsi dans le cadre d'analyse concernant les paramètres du parc immobilier influençant la consommation énergétique, ces corrélations devraient être prises en compte.

Un autre exemple traite de l'occupant car si on souhaite évaluer l'influence de l'occupation sur la performance énergétique, les charges internes dégagées par l'occupant, ainsi que le taux d'occupation vont faire partie des paramètres incertains. Lors de l'échantillonnage le lien logique entre présence de l'occupant et dégagement de charge interne supplémentaire dans le bâti devra être pris en compte pour être cohérent avec la réalité, mais cette prise en compte empêche la décomposition de Sobol pour entrées indépendantes. De la même façon, les problématiques liées aux conditions météorologiques doivent intégrer les corrélations entre chaque variable météo. Dans le cas de l'occupation et du climat, une difficulté supplémentaire s'ajoute de par leur caractéristique temporelle. Les corrélations sont doubles. Tout d'abord, l'aspect que nous avons illustré ci-dessus : la corrélation entre paramètres, que nous appellerons inter-corrélation. Puis la corrélation associée au caractère temporelle du signal : la corrélation du signal avec lui même dit auto-corrélation. Une entrée auto-corrélée est nommée entrée fonctionnelle. Les entrées spatiales et temporelles font partie de ce type d'entrée.

Ces deux aspects, climat et occupation, sont des enjeux clés actuels des bâtiments basse consommation. Ils ont une interaction forte avec le bâtiment puisqu'il s'agit des conditions limites du système : les sollicitations. Ces sollicitations proviennent de phénomènes complexes multi-physiques, que ce soit pour le métabolisme ou pour la climatologie. Ils sont stochastiques et leur influence au sein de bâtiment basse consommation d'énergie évolue. C'est donc en toute logique que de nombreux travaux récents en thermique du bâtiment traitent des incertitudes liées aux sollicitations. Par exemple, la caractérisation des scénarios d'occupation ainsi que leurs incertitudes traitée entre autres par Tanimoto et al. [64], Haldi et al. [21] et Parys et al. [48].

Dans la suite de ce manuscrit, seule la problématique des incertitudes associées aux entrées météorologiques est considérée. Nous verrons toutefois que la démarche développée est applicable à la gestion de l'occupation.

### **2.3.2 Incertitudes associées aux fichiers météorologiques**

Les fichiers météorologiques utilisés dans la simulation thermique dynamique sont des fichiers horaires. On distingue trois types de fichiers avec pour chaque, une utilisation bien définie :

- Les fichiers moyens annuels qui représentent la météo typique d'un site à partir de 20 années ou plus d'observations. Ce type de fichier est explicité en chapitre 3. Ce

sont les fichiers qui sont les plus utilisés, ils permettent d'obtenir une moyenne de la performance énergétique sur une année représentative du climat à long terme. Il en existe plusieurs formats.

- Les fichiers réels provenant d'observations brutes d'une année particulière. L'utilisation est liée aux problématiques de validation de modèle pour les comparaisons mesure/simulation.
- Les fichiers extrêmes qui permettent d'évaluer la performance d'un bâtiment en condition extrême. Il peut s'agir de fichiers construits à partir d'observations ou d'une année entière particulièrement extrême. Ils sont utilisés pour le dimensionnement de système énergétique.

En thermique du bâtiment, différents types d'incertitudes sont associées aux entrées météorologiques comme celles liées à la variabilité même de la météo, à la construction et au format des fichiers employés en simulation thermique dynamique ou bien des modèles sous-jacents de climatologie. Certains travaux modifient les fichiers moyens en fonction de l'évolution spatiale et temporelle du climat. Il peut s'agir de la volonté de prendre en compte le changement climatique (Jentsch et al. [30], [29]) ou bien des effets de micro-climat dus notamment au phénomène d'augmentation de la température dans les centres urbains que l'on appelle l'îlot de chaleur urbain (Chan et al. [6]). A titre d'illustration on peut énumérer des travaux portant sur l'évaluation des performances énergétiques de bâtiment dans le cas de :

- la variabilité météo (Wang et al. [68]);
- les périodes extrêmes (Breesch et al. [5]);
- l'îlot urbain (Kolokotroni et al. [32]);
- le changement climatique (Patidar et al. [49], Nik et al. [43] [42], DeWilde et al. [14], Tian et al. [66]);
- le type de format pour la simulation thermique dynamique (Crawley [11]);
- la résolution spatiale des fichiers météo utilisés en simulation thermique dynamique (Eames et al. [15]);
- le choix des données météo pour la calibration (Bhandari et al. [3]).

Ce tour d'horizon non exhaustif illustre l'engouement de la communauté sur les problématiques des incertitudes liées à la météo, et toutes exposent l'impact significatif de la modification de la météo sur le bâtiment. Ce type de travaux se base sur la propagation de différents échantillons pour visualiser la différence de comportement du bâtiment.

En fonction des études le nombre d'échantillons peut être compris entre quelques uns et jusqu'à plusieurs milliers. Par exemple, Nik et al. [42] réduisent une base de 771120 simulations par les méthodes d'échantillonnage de Monte-Carlo permettant une sélection dans l'espace des paramètres du parc immobilier suédois (152 bâtiments) et des échantillons de météo pour le changement climatique (douze scénarios de 140 ans chacun).

Il s'agit soit d'analyse d'incertitude et le résultat est la dispersion en sortie due aux incertitudes liées à la météo, soit d'étudier plus finement le comportement du bâtiment en fonction d'un nombre limité de scénarios. Dans ce dernier cas certaines études estiment les indices de sensibilité des entrées statiques sous différentes sollicitations météorologiques. On peut visualiser cette approche comme le fait d'évaluer le modèle à différents points de l'espace des paramètres (OAT).

A notre connaissance la quantification par analyse de sensibilité globale de l'influence de données temporelles est un sujet inédit en thermique du bâtiment. Ce manque s'explique en majeure partie par deux aspects :

- la prise en compte des corrélations en analyse de sensibilité
- l'échantillonnage cohérent et représentatif d'entrées temporelles

C'est également un challenge pour d'autres communautés qui disposent d'un contexte similaire de modèles complexes dynamiques à haute dimension nécessitant des entrées fonctionnelles comme la modélisation environnementale qui fait intervenir des entrées spatiales et parfois même temporelles. Voici des exemples d'approches développées dans la littérature.

### **2.3.3 Un besoin de méthodologie qui dépasse le domaine de la thermique du bâtiment**

Une première approche consiste à résumer l'information contenue dans la variable d'entrée fonctionnelle en un nombre limité de variables scalaires indépendantes. Confalonieri applique cette approche pour des modèles agro-climatiques [10]. L'auteur résume les données météorologiques en entrée du modèle aux entrées scalaires d'un générateur de météo. Dans ce genre d'approche la difficulté réside dans l'interprétation des résultats de sensibilité : plusieurs variables scalaires sont associées à une même variable fonctionnelle. De plus la caractérisation des incertitudes associées aux entrées scalaires n'est pas évidente car elle doit être cohérente avec l'incertitude de l'entrée fonctionnelle.

Une autre approche, développée par Lilburne et al. [36], évalue par groupe l'influence de données temporelles ou spatiales. Elle consiste à gérer l'ensemble des entrées corrélées comme un groupe de paramètres. Les méthodes basées sur la variance permettent en effet

d'évaluer la part de variance de la sortie  $Y$  induite par un groupe de variables d'entrées [59]. Pour pouvoir appliquer les méthodes de régression et FAST, les échantillons fonctionnels doivent être classés avant de leur associer un scalaire. Dans le cas des travaux de Saint-Geours et al. [54], des scalaires sont associés à des cartes permettant de traiter des problématiques de risque d'inondation. L'échantillonnage se fait en tirant aléatoirement parmi les scalaires sans remise, et donc parmi les cartes, ce qui suppose un nombre suffisant d'échantillons qui doivent être représentatifs des incertitudes de l'entrée. Les auteurs supposent également l'indépendance de l'entrée fonctionnelle avec les autres (qui peuvent être d'autres entrées fonctionnelles ou scalaires). La contrainte est de limiter l'échantillon, ce qui induit un certain biais, mais la méthode n'est pas adaptée pour évaluer la sensibilité de chaque entrée fonctionnelle dans le cas où elles seraient corrélées entre elles. Ainsi, dans le cadre de modèle agro-climatique lié à l'état hydrique des vignes, Guaus et al. [19] établissent l'influence des données météo en considérant comme un seul groupe un jeu de données climatiques composé de 4 variables météo corrélées. Les 22 séries météorologiques proviennent de mesure à différentes années et lieux de la zone étudiée.

Ainsi pour pouvoir utiliser les méthodes d'analyse de sensibilité standard, les entrées ou groupes d'entrées doivent être indépendants entre eux dans l'échantillonnage. Or, pour pouvoir être représentatif d'entrées fonctionnelles corrélées comme les variables météo, la corrélation entre les entrées doit être présente dans l'échantillonnage et l'incertitude doit être représentative de la physique du phénomène. Nous souhaitons obtenir l'influence de chaque entrée fonctionnelle en jeu. Nous proposons une méthodologie permettant d'échantillonner en respectant ces deux aspects. Pour cela on combine l'analyse par groupe pour s'affranchir de la corrélation temporelle. On aura l'effet global de la variable météo sur la plage de temps de la simulation. Pour la gestion de la corrélation entre variables, une méthode d'échantillonnage a été développée et est présentée en chapitre 3. Le principe est de générer un bruit aléatoire auto-corrélé auquel on rajoute une structure corrélative constante.



# Chapitre 3

## Méthodologie pour la prise en compte des données météorologiques dans les analyses de sensibilité

---

Dans ce chapitre, on décrit la méthode de génération des entrées fonctionnelles corrélées développée pour contourner le verrou scientifique de leur prise en compte en analyse de sensibilité. Cette méthodologie est ensuite appliquée dans le cadre de l'échantillonnage de la variabilité des entrées météo pour la simulation thermique dynamique. Ensuite, on vérifie les échantillons météo en testant leur indépendance d'échantillonnage et leur cohérence. Enfin les valeurs moyennes mensuelles sont évaluées et comparées avec des observations réelles afin de vérifier leur représentativité de la variabilité naturelle.

---

### 3.1 Méthode de génération développée

#### 3.1.1 Principe de la méthode

Afin de réaliser une analyse de sensibilité en prenant en compte les paramètres météorologiques, on doit échantillonner chaque variable. Par cela on entend générer pour chaque simulation des variations sur les données météorologiques de manière à induire une variance pour chaque entrée. Ces variations doivent être cohérentes avec les données climatiques du site mais doivent être également compatibles avec les méthodes d'analyse de sensibilité standard. Une méthodologie a été développée dans ce but et est décrite ci-après.

On considère un processus aléatoire que l'on note  $x_i(\theta, t)$  où  $t$  représente le temps. A noter que dans le cadre de problématique avec des entrées spatiales, la variable  $t$  peut être assimilée comme telle. La variable aléatoire  $\theta$  représente le caractère stochastique de la fonction. On peut définir ce processus aléatoire par une moyenne  $\bar{x}_i(t)$  à laquelle on rajoute une variable aléatoire  $\varepsilon_i(\theta, t)$  telle que :

$$x_i(\theta, t) = \bar{x}_i(t) + \varepsilon_i(\theta, t). \quad (3.1)$$

On suppose que la moyenne est déterministe alors que la variable aléatoire est définie par une fonction de covariance (ou auto-corrélation)  $C_{ii}(t, t')$  et une fonction de répartition  $h_i(t)$ . La fonction d'auto-corrélation représente la dépendance du signal avec lui-même (Ex : la température à 9 h du matin est influencée par la température à 8 h du matin). Comme  $\bar{x}_i$  est déterministe, la variabilité de  $x_i$  est due à  $\varepsilon_i$ .  $C_{ii}(t, t')$  est supposée décroissante et fonction de  $|t - t'|$ . On fait l'hypothèse dans le travail suivant que cette fonction peut être approximée par une fonction d'auto-corrélation décroissante. On suppose que la dépendance, par exemple, entre 8 h et 9 h est plus forte qu'entre 8 h et 12 h.

On peut extraire  $C_{ii}(t, t')$  et  $h_i(t)$  à partir d'une séquence météorologique donnée. Les fichiers météo considérés sont disponibles au pas de temps horaire. A partir du fichier météo source sélectionné, on peut évaluer les caractéristiques statistiques d'une journée type à partir du mois sélectionné. On estime la moyenne horaire, la fonction d'auto-corrélation  $C_{ii}(t_j, t_k)$  et on extrait la fonction de répartition empirique pour chaque pas de temps horaire  $h_i(t_j)$ ,  $t_j = 1, \dots, N_t$ . Comme la moyenne horaire est fixée, la génération d'une séquence cohérente dépend de la génération de la variable aléatoire  $\varepsilon_i(\theta, t)$ . Pour cela, on propose d'utiliser la procédure de Iman and Conover [27] ou bien celle de Karhunen-Loève qui sont toutes deux détaillées ci-après.

### 3.1.1.1 Génération du champ aléatoire $\varepsilon_i(\theta, t)$ par Iman et Conover

On note  $\mathbf{Z}^{(1)}$  un échantillon  $N \times N_t$  de valeurs normalement, identiquement et indépendamment distribuées et par  $\mathbf{X}$  un échantillon de la taille  $N \times N_t$  de variables indépendantes aléatoires distribuées selon  $\mathbf{h}_i$ . La procédure d'Iman et Conover [27] commence par générer un échantillon normal et corrélé  $\mathbf{Z}^c$  à partir d'une transformation de Cholesky :

$$\mathbf{C}_{ii} = \mathbf{U}_i^T \mathbf{U}_i, \quad (3.2)$$

$$\mathbf{Z}^c = \mathbf{Z}^{(1)} \mathbf{U}_i, \quad (3.3)$$

$\mathbf{U}_i$  est la matrice supérieure de Cholesky et l'exposant  $T$  est l'opérateur transposé. Puis, on note  $\mathbf{R}_z$  la matrice de corrélation des rangs associée à  $\mathbf{Z}^c$ , l'échantillon recherché est obtenu en réarrangeant l'échantillon  $\mathbf{X}$  selon  $\mathbf{R}_z$  de sorte que  $\mathbf{Z}^c$  et  $\mathbf{X}$  aient la matrice de corrélation des rangs (dont les éléments sont les coefficients de corrélation de Spearman),

$$\varepsilon_i = \mathbf{R}_z(\mathbf{X}).$$

Chaque ligne de  $\varepsilon_i$  fournit une séquence pour  $x_i(t_j)$ ,  $j = 1, \dots, N_t$  conformément à l'équation 3.1. Cette procédure est répétée pour chaque variable météorologique (dont l'indice associé est  $i$ ). Finalement, si nécessaire, une procédure similaire peut être employée pour imposer une structure corrélative entre les entrées météorologiques. Pour permettre l'application des méthodes standard d'analyse de sensibilité, nous ne réalisons pas cette dernière étape.

Nous avons expliqué dans la section 2.2.3.1 que deux échantillons étaient nécessaires pour obtenir, par une procédure de simple permutation, un jeu des indices de premier ordre (effet principal  $S_i$ ). Ces échantillons peuvent être construits en répétant la procédure de Iman et Conover avec un échantillon  $\mathbf{Z}^{(2)}$  différent mais avec le même échantillon  $\mathbf{X}$ .

### 3.1.1.2 Génération du champ aléatoire $\varepsilon_i(\theta, t)$ par Karhunen-Loève

Le champ aléatoire  $\varepsilon_i(\theta, t)$  peut être approximé en utilisant une décomposition de Karhunen-Loève [26] telle que le processus aléatoire s'écrive :

$$x_i(\theta, t) \sim \bar{x}_i + \sum_{ki=1}^{M_i} \sqrt{\lambda_{ki}} \xi_{ki}(\theta) f_{ki}(t) \quad (3.4)$$

avec  $\lambda_{ki}$  et  $f_{ki}$  respectivement les valeurs et fonctions propres de la fonction de covariance  $C_{ii}(t_j, t_k)$ ,  $\xi_{ki}(\theta)$  est une variable aléatoire indépendante et distribuée selon une loi normale centrée réduite.  $M_i$  est le nombre de termes (ou modes) de Karhunen-Loève (KL). KL se base sur la décomposition en  $M_i$  fonctions et valeurs propres de la fonction de covariance. Le nombre  $M_i$  de termes (ou modes) est, en pratique, fixé aux  $M_i$  premiers termes contenant 95% de la variance du processus aléatoire  $x_i(t)$ .

La difficulté pour la génération de champs aléatoires avec KL réside dans la capacité à déterminer précisément les valeurs et fonctions propres. La méthode des ondelettes de Galerkin [50] permet de simplifier ce calcul.

Le principe de KL rejoint l'approche de gestion des entrées fonctionnelles décrites en section 2.3.3, qui est de résumer le signal temporel (ou spatial) en un nombre fini de scalaires. A chaque entrée est associée un nombre  $M_i$  de termes. La sensibilité de l'entrée fonctionnelle est déterminée à partir du groupe de scalaires  $M_i$  associé.

L'approche de génération du champ aléatoire par KL a été mise en œuvre dans le cadre du projet CNRS PEPS<sup>1</sup> ASenDyn (Analyse de Sensibilité globale pour les modèles

<sup>1</sup>Projets Exploratoires Premier Soutien

Dynamiques) [1]. La méthode Iman et Conover est celle qui est utilisée dans le cadre de cette thèse.

La section suivante applique la méthode de génération dans un cas concret des données météorologiques d'un site pour la STD.

## 3.2 Application de la méthode sur des conditions climatiques

### 3.2.1 Conditions initiales de la génération

Afin d'appliquer la méthodologie décrite, nous devons tout d'abord extraire les informations statistiques nécessaires : la fonction d'auto-corrélation, la loi de distribution et la moyenne. Pour être compatible avec la simulation thermique dynamique, les données météorologiques sont au pas de temps horaire. Nous souhaitons pouvoir, à partir de données sources, reconstruire des données météorologiques cohérentes. Pour ce faire, nous devons avoir un minimum d'échantillons statistiques. Il se pose alors la question de la disponibilité des observations météorologiques. En effet, des données mesurées au pas de temps horaire sur le long terme sont peu accessibles. Nous avons donc fait le choix de prendre comme fichier source des fichiers moyens développés et disponibles pour la STD.

#### 3.2.1.1 Le fichier source

Des données météorologiques sont nécessaires à la simulation thermique dynamique. Afin d'obtenir des résultats de performance d'un bâtiment qui soient représentatifs du lieu d'implantation, des fichiers moyens ont été créés par la communauté [11]. Ces fichiers représentatifs d'une durée d'un an et d'un pas de temps horaire sont des années typiques météorologiques ou TMY pour typical meteorological year. On les appelle également TRY pour test reference year ou bien des DRY pour design reference year. Pour constituer ces fichiers typiques, chaque mois est sélectionné dans une base de données météorologiques mesurées sur des dizaines d'années. Ainsi sur 20 ans d'observations météorologiques, le choix d'un mois de janvier typique va se faire parmi les 20 mois de janvier mesurés. Ce choix se fait en fonction de la statistique extraite des 20 ans d'observations. Le mois le plus représentatif des mois de janvier sur 20 ans est retenu. Une des méthodes de sélection les plus communes se base sur la méthode statistique de Filkenstein-Schafer [17]. Il s'agit d'une approche empirique qui sélectionne des mois individuels sur plusieurs années de mesures.

Douze mois typiques sont choisis individuellement parmi les années d'observations,

puis ces douze mois sont concaténés en une année. Les paramètres qui servent de base pour la sélection sont des indices calculés à partir des valeurs horaires de la température sèche, du point de rosée, de l'humidité relative, de la vitesse du vent et de l'intensité du rayonnement solaire global. La statistique de chaque mois mesuré est calculée à partir de ces indices. Le mois qui a la somme des indices la plus proche des indices du mois typique est sélectionné. Ces indices sont journaliers et sont calculés pour chaque mois de chaque année.

Les TMY sont disponibles pour de nombreuses villes aux Etats-Unis mais sont plus rares pour le reste du monde. Pour pallier à ce manque, un programme de l'ASHRAE (American Society of Heating, Refrigerating and Air-conditioning Engineers) [13] développe les TMY pour d'autres localisations, il s'agit des IWEC pour international weather energy calculations. Ces fichiers sont mis à disposition au format du logiciel EnergyPlus [45]. Pour la France, douze localisations sont disponibles dont Lyon et Strasbourg. La période d'observation sur laquelle les fichiers sont construits comprend les années de 1982 à 1999. Il s'agit du type de fichier utilisé à titre d'exemple dans ce travail de thèse. Ce choix permet une reproductibilité des travaux car ils se basent sur des fichiers en libre accès.

#### 3.2.1.2 La période source

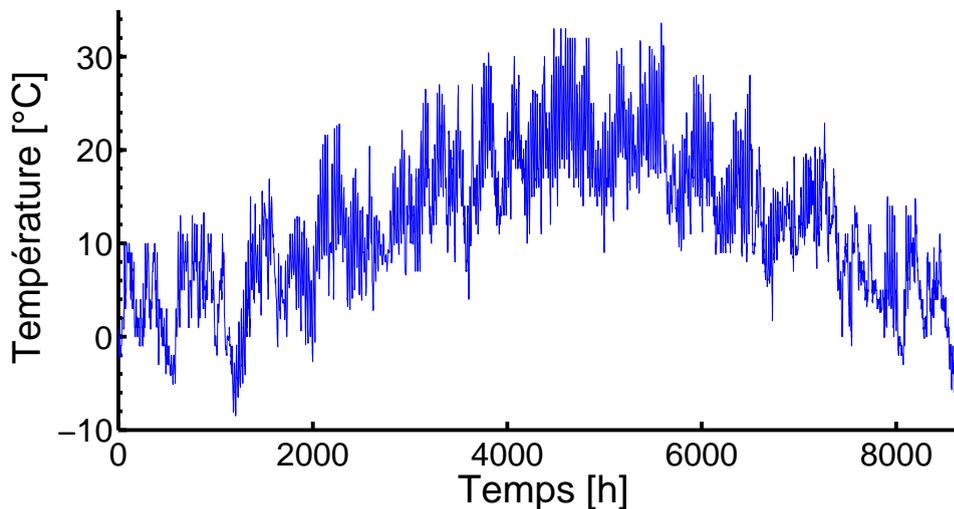
Les informations telles que la fonction d'auto-corrélation, la loi de distribution ainsi que la moyenne doivent être obtenues à partir d'un échantillon statistique. Dans le cas où plusieurs années mesurées sont disponibles la variabilité statistique est évidente. Or nous souhaitons extraire les informations statistiques des fichiers moyens tel que le IWEC. Ces fichiers horaires ne sont que d'une durée d'un an. Pour obtenir les grandeurs statistiques nous proposons de les extraire à partir de la variabilité de la journée sur un mois donné. Pour cela nous faisons l'hypothèse qu'un mois est la pseudo-répétition de journées aléatoires par translation.

Considérons un mois de 31 jours, si on concatène les 744 valeurs horaires en matrice de  $31 \times 24$  on a :

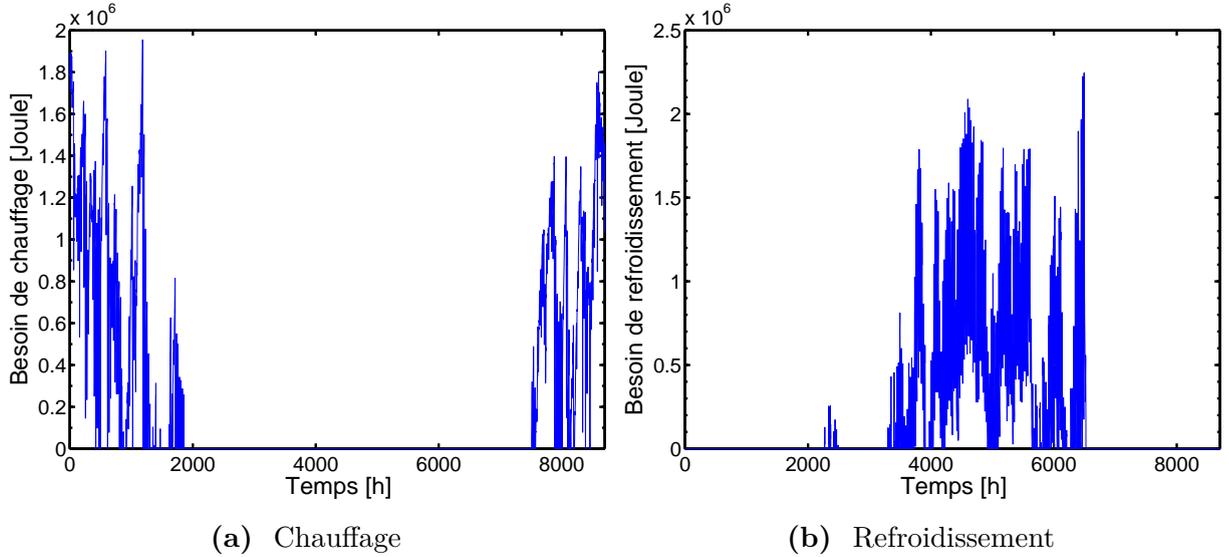
$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{744} \end{bmatrix} = \begin{bmatrix} a_{1,1} & \dots & a_{1,24} \\ a_{2,1} & \dots & a_{2,24} \\ \vdots & \dots & \vdots \\ a_{31,1} & \dots & a_{31,24} \end{bmatrix} \quad (3.5)$$

On extrait ensuite la moyenne horaire, l'auto-corrélation sur la journée ainsi que la distribution cumulée à chaque pas de temps horaire, soit 24 distributions cumulées empiriques. Cette approche ne peut fonctionner qu'avec les mois de l'année les plus proches de

la stationnarité. Par cela on entend que le mois soit sans forte tendance saisonnière. Cette hypothèse permet de considérer dans le mois donné qu'il est équiprobable d'obtenir une journée ou une autre. Dans le cadre du mois de mars cette hypothèse n'est pas possible car il y a une tendance importante d'augmentation de température. Les inter-saisons ne sont pas envisageables. Les deux périodes stables de l'année sont l'hiver et l'été. Comme nous pouvons le voir sur la figure 3.1 qui représente l'évolution de la température extérieure sur l'année à Lyon. Entre les deux phases de croissance et de décroissance de température durant l'année, se situe deux "plateaux". On ne considère dans notre étude que le mois de janvier et de juillet, les mois centraux de chaque saison. Les périodes choisies sont en adéquation avec les périodes intéressantes à explorer pour évaluer le comportement de la maison INCAS-DM, il s'agit des périodes où les besoins de chauffage et de refroidissement sont les plus importants. La figure 3.2 représente les besoins horaires à gauche de chauffage et à droite de refroidissement pour le cas d'une simulation annuelle de la maison INCAS-DM. On observe que les besoins sont majoritaires sur la période hivernale et également présents sur la période estivale ce qui est due notamment aux faibles infiltrations et à l'absence de renouvellement d'air dans le modèle thermique.



**Figure 3.1** – Évolution de la température du fichier moyen (IWEC) de Lyon. On observe deux phases stables qui correspondent à l'hiver et l'été.



**Figure 3.2** – Graphes des besoins horaires sur l’année a) de chauffage et b) de refroidissement pour le climat de Lyon de la maison INCAS-DM. Les besoins principaux se situent sur les périodes hivernales et estivales

Nous illustrons l’extraction des données statistiques pour la ville de Lyon en juillet à partir du fichier IWEC.

### 3.2.2 Extraction des données statistiques

L’étape de l’extraction consiste à collecter les informations nécessaires à la génération des échantillons. On extrait ces informations pour chaque entrée météo à partir du jeu de données du fichier IWEC source (i.e. un mois). Pour rappel les informations qui caractérisent les données météo sont : l’auto-corrélation, la fonction de répartition, et la moyenne horaire journalière. A partir de ces informations il est possible de générer des journées aléatoires typiques de la période sélectionnée (ici janvier et juillet) et de construire des mois cohérents au regard de cette période. Le procédé d’extraction se déroule comme suit pour chaque entrée météo :

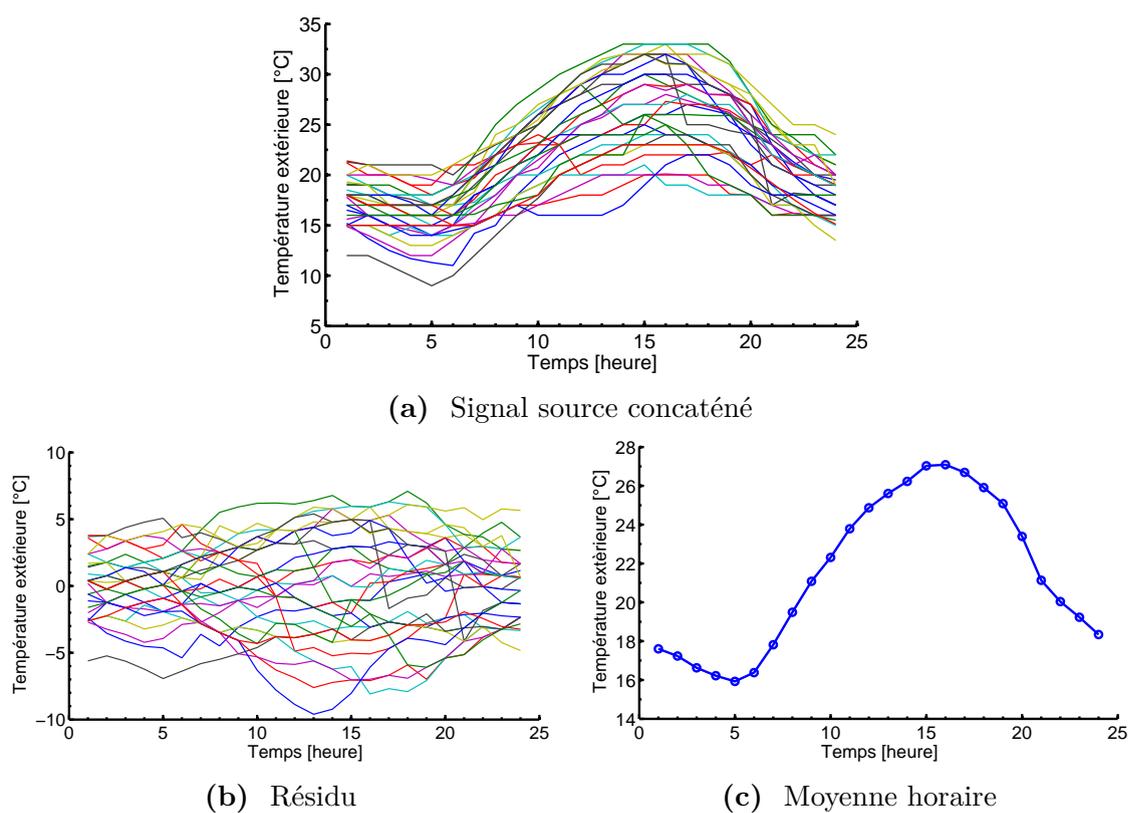
1. Les données horaires des 31 jours de janvier ou juillet sont rassemblées dans une matrice  $31 \times 24$ ,  $D_i(d_n, t_j)$ .
2. On calcule la moyenne horaire  $\bar{x}_i(t_j)$ ,  $t_j = 1, \dots, 24$  et la variation horaire une fois la moyenne extraite  $r_i(d_n, t_j) = D_i(d_n, t_j) - \bar{x}_i(t_j)$ .
3. La matrice d’auto-covariance est estimée de la manière suivante :  $C_{ii}(t_j, t_k) = \frac{1}{N-1} \sum_{d_n=1}^{31} r_i(d_n, t_j) \times r_i(d_n, t_k)$ . Elle peut être normalisée afin d’obtenir une fonction

d'auto-corrélation. On peut aussi approximer l'auto-corrélation par une fonction décroissante (souvent une exponentielle).

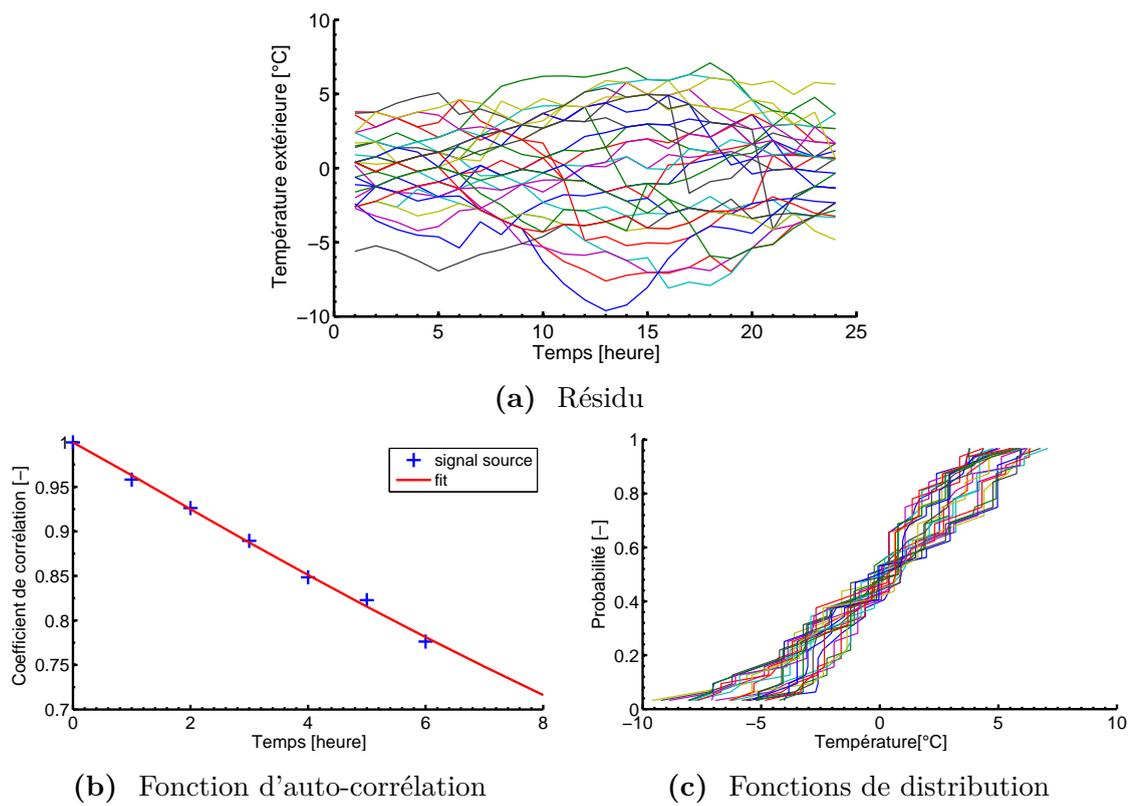
4. La fonction de répartition horaire  $h_i(t_j)$  empirique est extraite pour chaque pas de temps.

Pour illustrer les étapes ci-dessus, l'exemple de la température extérieure du mois de juillet du fichier IWEC de Lyon (station de l'aéroport St-Exupéry) est présenté. Les auto-corrélations des variables température, vitesse et direction de vent et humidité relative sont estimées la nuit pour éviter les effets d'interactions avec le rayonnement solaire. Dans le cas du rayonnement solaire direct normal et diffus horizontal, les informations statistiques ne sont considérées que sur la période diurne.

La figure 3.3 montre la première étape de l'extraction. On dispose de 31 valeurs pour chaque pas de temps  $t$  du mois choisi. La figure 3.3a représente les 31 jours tracés pour la température du mois de juillet pour Lyon. De cette dispersion horaire est extraite la moyenne horaire (figure 3.3c). On obtient le résidu du signal (figure 3.3b). De ce résidu est extrait la fonction d'auto-corrélation (figure 3.4b) et les fonctions de distribution horaires (figure 3.4c).



**Figure 3.3** – Procédure d'extraction à partir du signal concaténé  $31 \times 24$  représenté en a). On extrait en b) le signal sans la moyenne horaire (le résidu) et en c) la moyenne horaire



**Figure 3.4** – Procédure d'extraction à partir des valeurs horaires sans la moyenne en a), dont sont extraits en b) la fonction d'auto-corrélation et en c) les fonctions de distribution horaires empiriques

### 3.2.3 Génération des échantillons

Quand les données statistiques sont extraites, la génération des échantillons peut être faite. Pour  $N$  jeux de  $p$  entrées d'une durée de  $N_t$  les 4 étapes de la génération sont les suivantes :

1. la génération de  $N \times N_t$  échantillons aléatoires et indépendants et ce pour les  $p$  entrées
2. l'ajout de l'auto-corrélation selon  $t$  aux  $p \times N$  échantillons
3. la transformation de la loi normale initiale en loi de distribution empirique pour chaque pas de temps horaire
4. l'ajout de la moyenne déterministe

On applique l'extraction et la génération sur 6 entrées, les prédominantes en météorologie et nécessaires à la plupart des logiciels de simulation thermique dynamique : la température sèche extérieure, le rayonnement solaire direct normal et diffus horizontal, l'humidité relative, la vitesse et la direction de vent. Les figures représentant les informations statistiques clés pour chaque variable météo sont regroupés en annexe.

On génère 1000 mois de juillet pour Lyon. A titre d'exemple, la figure 3.5 représente 20 échantillons tirés aléatoirement parmi les 1000. Le comportement de ces échantillons semble adéquat. Nous n'observons pas de valeurs trop écartées du signal de référence provenant du fichier IWECC. Comme la génération se fait à partir des fonctions de distribution horaires empiriques, les seuils limites de chaque variable sont respectés correspondant aux limites physiques associées à la variable comme l'humidité relative qui ne peut pas être supérieure à 100% ou bien le rayonnement solaire qui est limité à une certaine intensité.

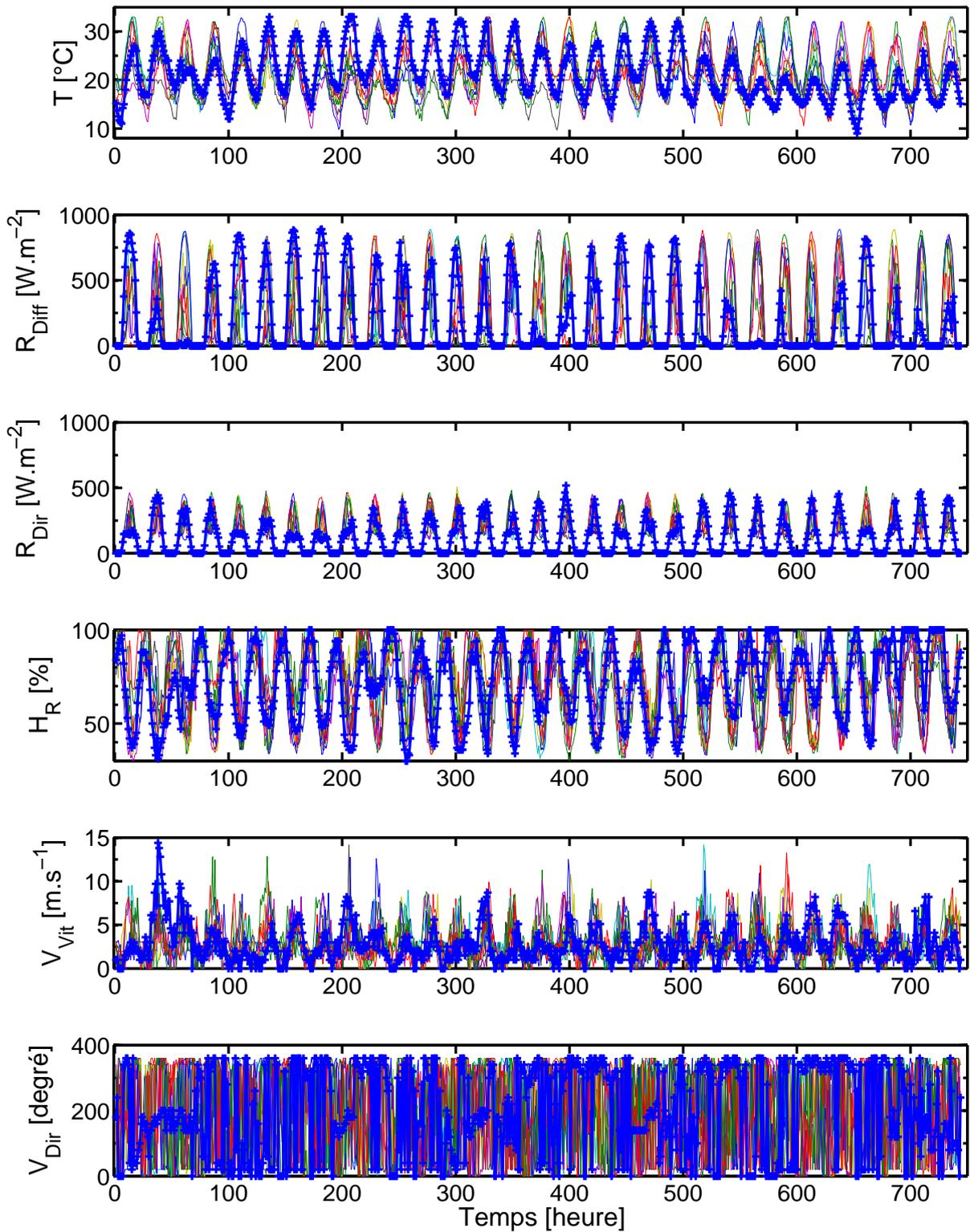


Figure 3.5 – Exemple de 20 échantillons obtenus parmi les 1000 pour les six variables. La courbe bleue est le signal source provenant du IWEC de Lyon pour le mois de juillet.

### 3.3 Vérification de la cohérence des échantillons

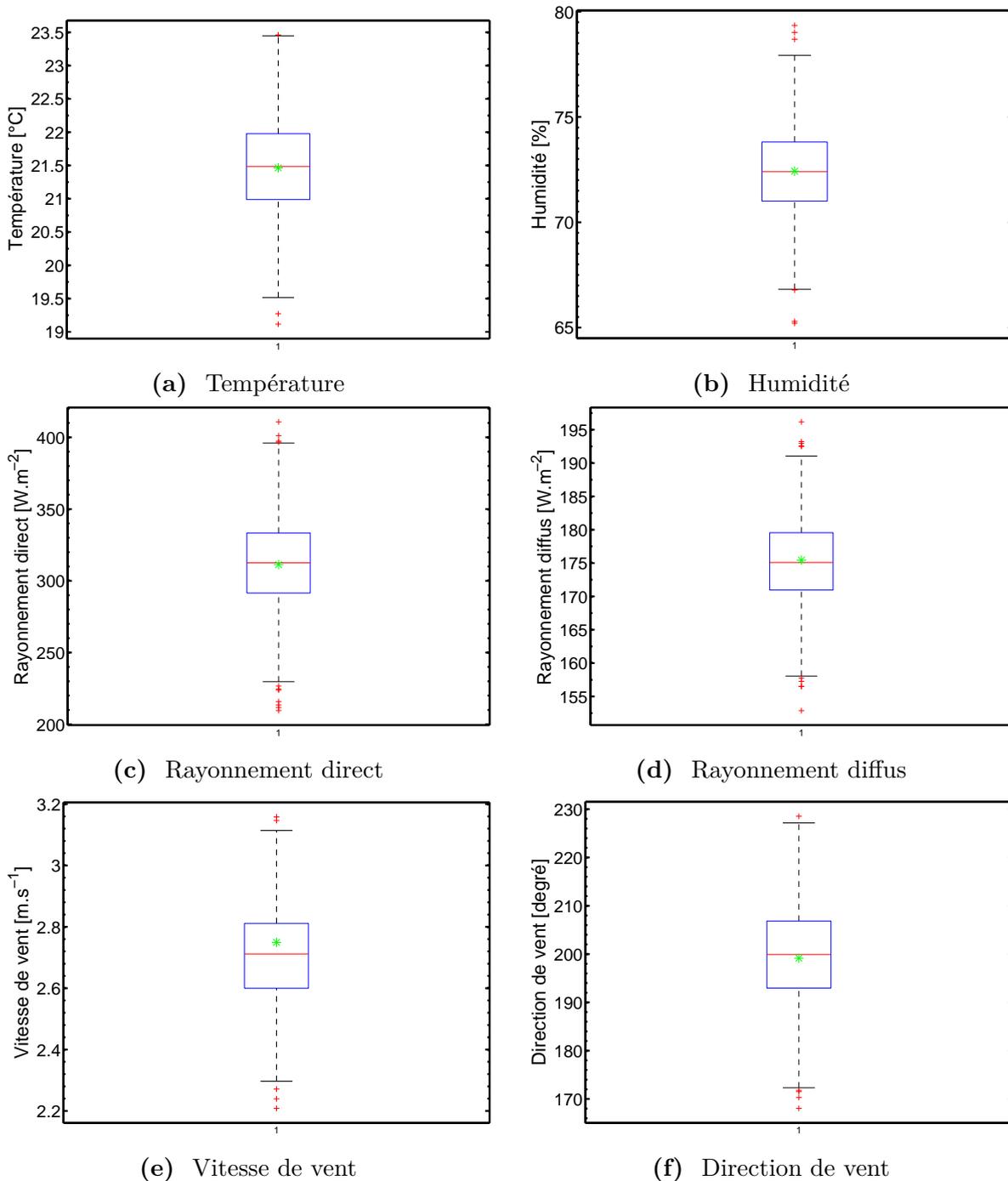
La vérification visuelle que l'on peut avoir sur les échantillons ne suffit évidemment pas, nous allons devoir tester leurs propriétés. Par plusieurs calculs et représentations graphiques on tente d'avoir un retour sur les échantillons que l'on a générés. Tout d'abord les fonctions d'auto-corrélation de sortie ainsi que les fonctions de distribution sont comparées entre références et échantillons afin de veiller à ce que les échantillons de sortie respectent bien les informations statistiques en amont de la génération.

Ces conditions remplies, d'autres critères vont permettre de vérifier les échantillons générés, comme la dispersion des valeurs par rapport au fichier source, mais aussi le recouvrement de l'espace des paramètres. Puis les corrélations entre variables le long du processus de génération est vérifié et enfin les ordres de grandeur de variations des échantillons sont comparés à des normales mensuelles sur dix ans de Météo-France.

#### 3.3.1 Dispersion et accord avec le fichier source

On vérifie tout d'abord le positionnement des données météorologiques de référence c'est à dire provenant du IWECC par rapport aux échantillons générés. Les informations sont représentées sous forme de boxplot. La représentation par boxplot ou boîte à moustache est utilisée pour visualiser la dispersion liée aux échantillons. Cette représentation permet de comparer des jeux de données différents sans hypothèse sur la symétrie ou la nature de la dispersion. La représentation se fait à l'aide de paramètres de localisation : la médiane (50% de l'information), le 1<sup>er</sup> ( $Q_1$ ) et le 3<sup>e</sup> quartile ( $Q_3$ ). Ces points de localisation sont moins influencés par des valeurs extrêmes que peut l'être la moyenne empirique. Pour la détection des valeurs extrêmes cette représentation est conseillée plutôt que de faire l'hypothèse forte d'une loi normale et des seuils de confiance par rapport au sigma. Les points atypiques sont repérés par rapport aux "moustaches". Les extrémités des moustaches sont délimitées par 1.5 fois l'intervalle inter-quartile ( $Q_3 - Q_1$ ), ce qui représente plus de 99% de l'information.

La moyenne sur le mois de la valeur moyenne journalière des 1000 mois échantillonnés est calculée pour chaque variable. La figure 3.6 représente les boxplots pour la dispersion des six variables. L'étoile verte représente la valeur calculée pour chaque variable du fichier source IWECC pour Lyon pour le mois de juillet. Cette figure 3.6 nous permet d'apprécier les dispersions générées sur des valeurs journalières mensuelle pour chaque mois généré. La valeur de référence est bien située proche de la médiane pour les six variables.



**Figure 3.6** – Dispersion des valeurs moyennes journalières pour les 1000 mois de l'échantillon et les six variables. L'étoile verte représente la valeur moyenne journalière pour le mois de juillet provenant du fichier IWEAC de Lyon.

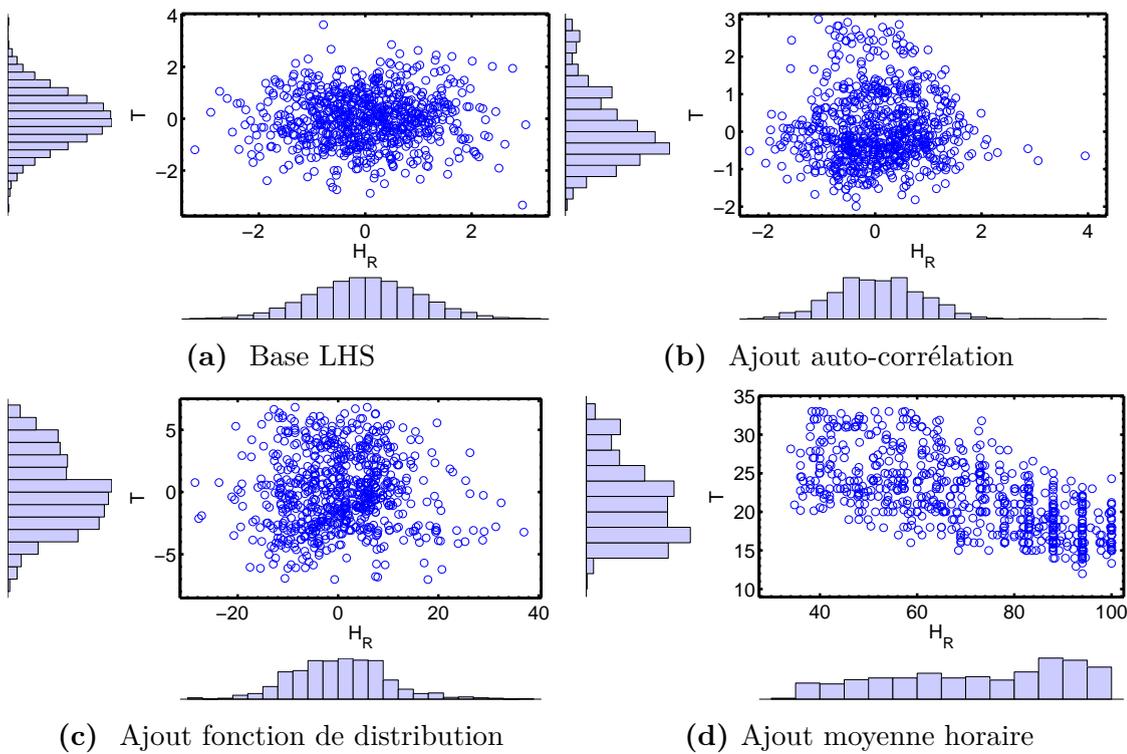
- a) Température journalière moyenne
- b) Humidité relative journalière moyenne
- c) Rayonnement solaire direct normal journalier moyen
- d) Rayonnement solaire diffus horizontal journalier moyen
- e) Vitesse de vent moyenne journalière
- f) Direction de vent moyenne journalière

### 3.3.2 Évolution de l'inter-corrélation durant le processus de génération

La base de la méthodologie développée réside dans la volonté de faire des analyses de sensibilité en prenant en compte les entrées fonctionnelles corrélées. Pour cela nous souhaitons utiliser des méthodes standard basées sur la décomposition de la variance, or ces méthodes nécessitent une hypothèse d'indépendance dans la génération des entrées. L'idée a été de générer de manière indépendante une variable aléatoire (ou bruit) à laquelle est rajoutée un signal déterministe constant (ici la moyenne horaire) pour chaque échantillon et propre à chaque variable. On suppose donc que seul l'ajout de ce signal déterministe induit une corrélation entre variables. Pour vérifier cette hypothèse, on évalue les inter-corrélations à chaque étape du processus de génération.

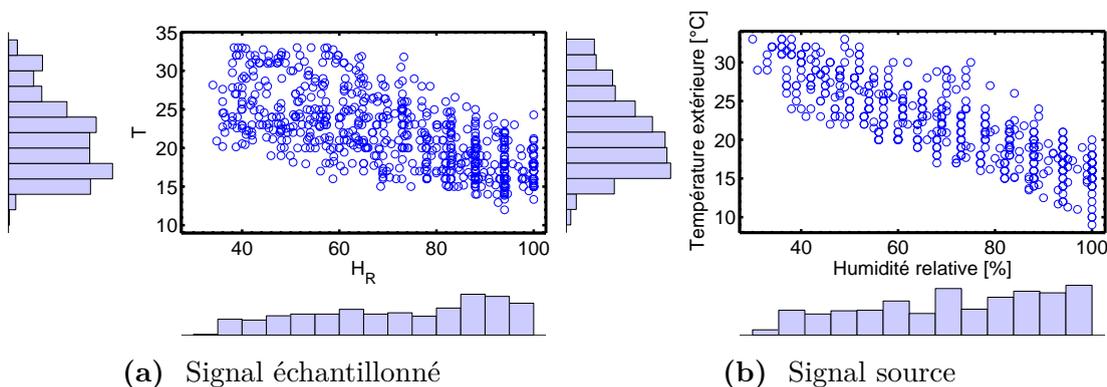
Nous proposons d'illustrer cette approche par la figure 3.7 qui représente la température en fonction de l'humidité sur les 744 valeurs d'un échantillon parmi les 1000. Cette représentation graphique dit scatterplot, permet de visualiser le lien entre variables, un précédent exemple de cette représentation a été présenté en figure 2.1 (réhabilitation). Un nuage de point sous forme de cercle témoigne d'une indépendance alors qu'une tendance témoigne d'une corrélation. A coté de chaque axe est représenté la distribution des données sous forme d'histogramme et nous pouvons observer son évolution durant les étapes de la génération :

- En figure 3.7a on observe le jeu de variable aléatoire de départ échantillonné selon LHS sur une loi normale centrée réduite. On visualise cette loi normale sur les histogrammes encadrant le graphe. Aucune structure corrélative n'est observée ce qui est confirmé par le calcul du coefficient de corrélation de Pearson  $\rho = 0,0308$ .
- En figure 3.7b chaque variable a été auto-corrélée selon la fonction d'auto-corrélation par la procédure de Iman et Conover. Cette transformation perturbe la loi normale de départ comme on peut l'observer. Aucune corrélation entre les variables n'est observée :  $\rho = 0,0196$ . Évidemment une auto-corrélation est présente mais n'induit pas d'inter-corrélation.
- En figure 3.7c chaque variable a sa distribution modifiée et adaptée à l'empirique. Pas de corrélation entre variable  $\rho = 0,0198$ . On observe bien une transformation des distributions entre 3.7b et 3.7c.
- En figure 3.7d la moyenne horaire extraite est rajoutée. On observe cette fois une forte corrélation entre variables, confirmée par le coefficient de corrélation  $\rho = -0,6790$ .



**Figure 3.7** – Corrélation entre la température et l’humidité aux différentes étapes de génération, pour un jeu d’échantillon (1 mois) tiré aléatoirement parmi les  $N$ . L’inter-corrélation n’est présente qu’après l’ajout du signal déterministe.

- a) Base de l’échantillonnage par LHS sur une loi normale centrée réduite
- b) Ajout à l’échantillon de base de l’auto-corrélation par Iman et Conover
- c) Ajout de la fonction de distribution empirique
- d) Ajout de la moyenne horaire



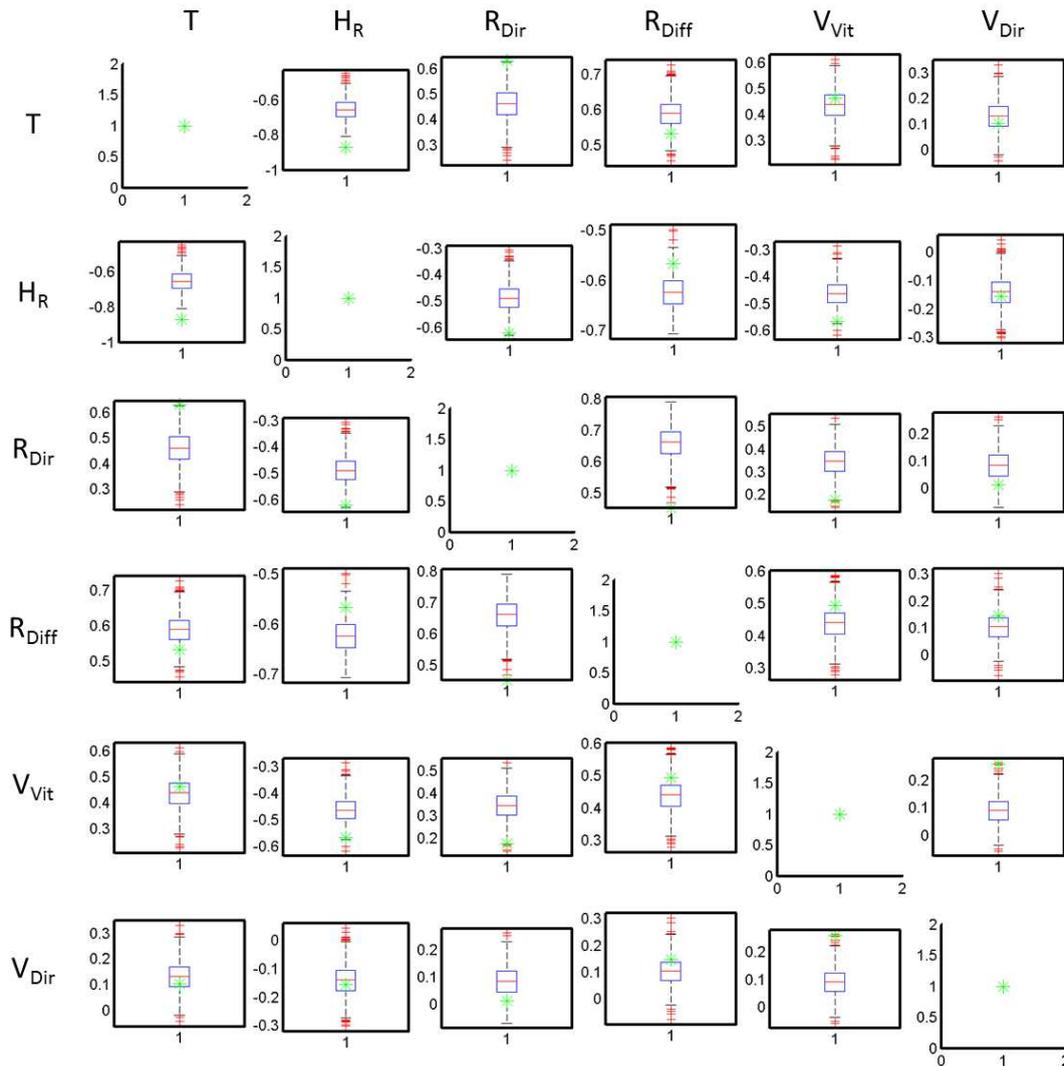
**Figure 3.8** – Comparaison de la corrélation entre  $T$  et  $H$  entre a) un échantillon tiré au hasard dans les  $N$  et b) le signal source. La structure corrélative est similaire

Pour vérifier la cohérence de l'inter-corrélation obtenue, nous comparons l'échantillon avec le signal source en figure 3.8. Nous visualisons une structure corrélative similaire.

Le choix de l'humidité et de la température comme illustration n'est pas anodin, il s'agit des variables les plus fortement corrélées dans notre signal source comme nous pouvons le constater dans le tableau 3.1 qui regroupe les valeurs d'inter-corrélations pour les six variables du signal source. La valeur de la corrélation entre humidité et température est de  $\rho = -0,8691$ . Si on compare cette valeur à celle obtenue précédemment, la corrélation entre l'humidité et la température est sous-estimée mais du même ordre de grandeur. Pour pouvoir évaluer si cette sous-estimation est ponctuelle, nous calculons les inter-corrélations des échantillons des six variables. La figure 3.9 regroupe la dispersion de ces inter-corrélations et se lit comme le tableau, c'est-à-dire comme une matrice d'inter-corrélation. L'étoile verte représente la valeur de référence. Nous observons que certaines inter-corrélations sont minimisées comme pour les corrélations formées par le triplet humidité, température et rayonnement direct ( $\rho_{T-H_R}$ ,  $\rho_{R_{Dir}-H_R}$ ,  $\rho_{R_{Dir}-T}$ ). D'autres corrélations sont maximisées, le décalage le plus élevé concerne la corrélation entre le rayonnement direct et le rayonnement diffus ( $\rho_{R_{Dir}-R_{diff}}$ ). Même dans le cas de décalage entre la dispersion des échantillons et la valeur de corrélation de référence, la tendance corrélative est respectée.

**Table 3.1** – Valeurs des coefficients de corrélation de Pearson des six variables pour le signal source

<b>Variables</b>	$T$	$H_R$	$R_{Dir}$	$R_{Diff}$	$V_{vit}$	$V_{dir}$
$T$	1,0000	-0,8691	0,6316	0,5323	0,4636	0,1020
$H_R$	-0,8691	1,0000	-0,6204	-0,5659	-0,5660	-0,1543
$R_{Dir}$	0,6316	-0,6204	1,0000	0,4527	0,1768	0,0129
$R_{Diff}$	0,5323	-0,5659	0,4527	1,0000	0,4932	0,1473
$V_{vit}$	0,4636	-0,5660	0,1768	0,4932	1,0000	0,2599
$V_{dir}$	0,1020	-0,1543	0,0129	0,1473	0,2599	1,0000



**Figure 3.9** – Représentation en matrice de corrélation entre les variables. A chaque  $\rho_{x_i-x_j}$  est associée une représentation en boxplot des valeurs d’inter-corrélations pour les 1000 échantillons générés, ainsi que la valeur du coefficient de corrélation du signal source (étoile verte) répertoriée dans le tableau 3.1. On observe certains décalages et particulièrement pour les forts coefficients de corrélation, les valeurs de coefficient restent toutefois cohérentes avec la physique sous-jacente

### 3.3.3 Couverture de l'espace des paramètres

Nous souhaitons désormais évaluer la capacité des échantillons à recouvrir l'espace des paramètres de la même manière que le signal IWEC source. La vérification consiste à tracer chaque variable en fonction d'une autre (scatterplot) pour les 744 valeurs du signal source ainsi que pour dix échantillons parmi 1000. Cette approche va nous permettre de visualiser les zones recouvertes dans l'espace des paramètres et si cela est en adéquation avec le signal source. Le recouvrement de l'espace des paramètres est lié aux corrélations entre variables, et nous avons vu précédemment qu'il existe des écarts d'inter-corrélation entre certaines variables.

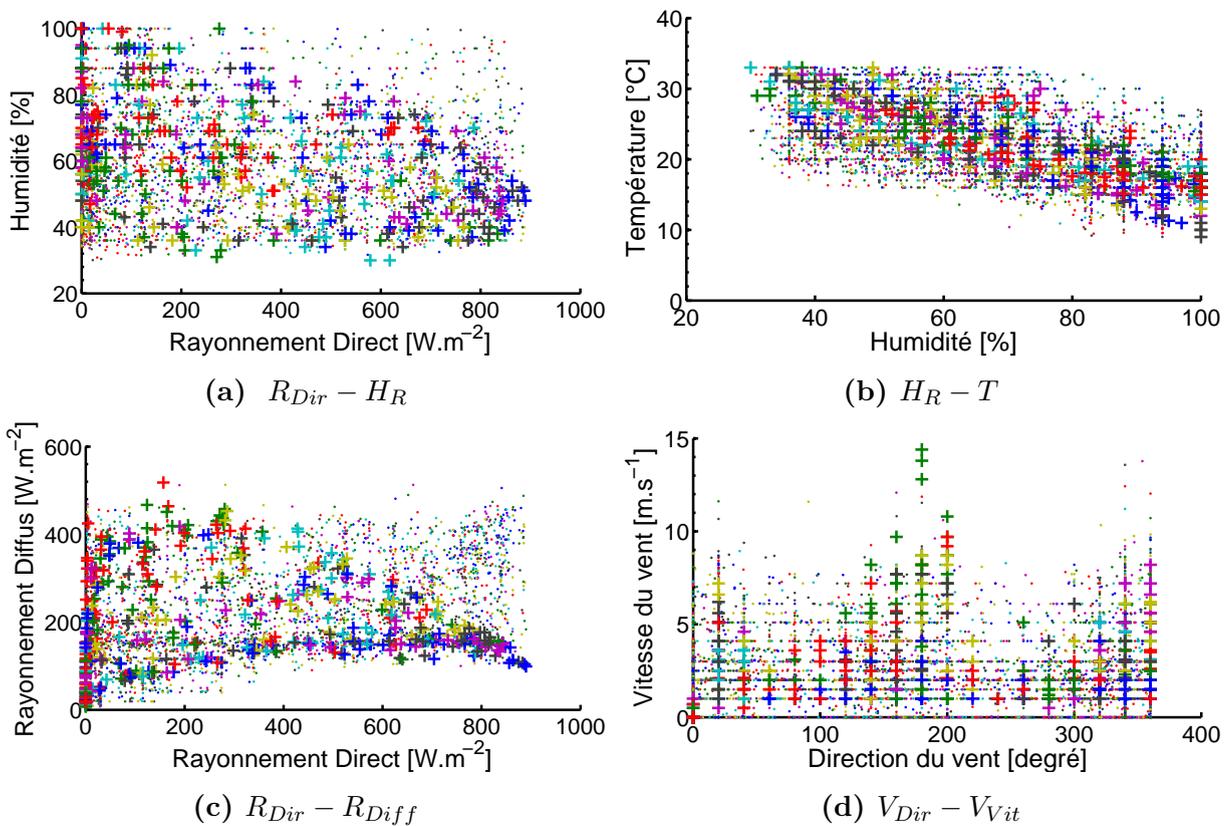
Parmi les quinze combinaisons de variables possibles, nous présentons en figure 3.10 quatre exemples de recouvrement de l'espace qui révèlent des disparités entre les échantillons et le signal source.

On visualise sur la figure 3.10 les scatterplot pour les couples  $R_{Dir} - H_R, H_R - T, R_{Dir} - R_{Diff}$  et  $V_{Dir} - V_{Vit}$ . On observe pour ces quatre graphes des zones de l'espace des paramètres couvertes par les échantillons et non par le signal source. Par exemple, les figures 3.10a et 3.10c montrent l'existence dans les échantillons de valeurs élevées respectivement d'humidité et de rayonnement diffus quand le rayonnement direct est élevé. Cela induit une surestimation des inter-corrélations des échantillons par rapport à celles du signal source comme nous l'avons constaté en figure 3.9. Ces artefacts sont le résultat de la perte d'information dans la méthode d'échantillonnage qui repose sur le fait que la corrélation entre les entrées s'exprime par la moyenne horaire déterministe. Cette hypothèse est nécessaire pour réaliser par la suite une analyse de sensibilité car elle suppose l'indépendance dans la génération des entrées.

Ainsi cette perte d'informations s'explique par les informations statistiques extraites qui ne prennent pas en compte toutes les dépendances entre variables, il existe en effet des phénomènes sous-jacents que nous ne reproduisons pas dans l'échantillonnage. Par exemple, nous observons en figure 3.10c des valeurs de rayonnement diffus échantillonnées qui sont élevées lorsque le rayonnement direct échantillonné est important ce qui n'est pas cohérent avec les phénomènes de météorologie. En effet, le rayonnement diffus est élevé lorsque la couverture nuageuse est importante contrairement au rayonnement direct qui est élevé lorsque le ciel est clair. La variable sous-jacente à la relation entre rayonnement diffus et direct pour les valeurs élevées est la nébulosité, c'est-à-dire la part de couverture nuageuse dans le ciel. Or dans la méthode d'échantillonnage la différence de comportement des rayonnements en fonction de la couverture nuageuse n'est pas prise en compte.

De la même manière, on observe en figure 3.10d que pour le signal source, la vitesse de vent est conditionnée par la direction de vent. Pour des valeurs de directions comprises entre  $150^\circ$  et  $200^\circ$  la vitesse de vent est plus élevée que pour des directions comprises entre  $200^\circ$  et  $250^\circ$ . L'échantillonnage prend plus difficilement en compte les directions préférentielles pour la vitesse de vent.

Ces constatations mettent l'accent sur la complexité des phénomènes météorologiques car de nombreuses variables sont en interaction et les inter-corrélations dépendent des valeurs de chaque variable qui peuvent avoir un effet "interrupteur" comme la présence ou non de nuages qui peut être assimilée à une variable aléatoire qui induit une différence de comportement entre les rayonnements. Cependant à part le rayonnement direct et diffus, les écarts de recouvrement de l'espace ne sont pas significatifs. Par ailleurs sachant que l'échantillon source comparatif est limité, les résultats sont satisfaisants car les corrélations, les valeurs limites et le comportement des variables sont respectés.



**Figure 3.10** – Comparaison du recouvrement de l'espace des paramètres par scatterplot de 10 échantillons tirés au hasard dans les 1000 (point) et du signal source (croix), pour des singularités.

- a) Humidité en fonction du rayonnement direct ( $R_{Dir} - H_R$ )
- b) Température en fonction de l'humidité ( $H_R - T$ )
- c) Rayonnement diffus en fonction du rayonnement direct ( $R_{Dir} - R_{Diff}$ )
- d) Vitesse du vent en fonction de sa direction ( $V_{Dir} - V_{Vit}$ )

### 3.3.4 Comparaison aux normales mensuelles de Météo-France

Les échantillons générés sont censés être représentatifs de la variabilité météorologique du site sélectionné. Ainsi, pour vérifier la cohérence des échantillons générés, l'idéal est de pouvoir les comparer à des données réelles. Cependant, il est difficile d'obtenir des observations réelles de qualité sur une durée excédant l'année. Dans le cadre de cette thèse la vérification de la cohérence des échantillons va donc se limiter à une comparaison aux normales mensuelles pour une durée de dix ans de 2004 à 2013. Il s'agit des données disponibles par Météo-France (MF)[40]. Il est également possible de récupérer des jeux de données provenant de stations d'amateurs, mais bien qu'ayant accès à des bases de données plus importantes, nous connaissons pas les incertitudes associées, nous nous fierons seulement aux données de Météo-France.

Par ailleurs il est important dans la comparaison de calculer les normales avec le même protocole que Météo-France qui provient de l'Organisation Mondiale de la Météorologie (OMM). Ces protocoles ont été mis en place afin de permettre une synchronisation des observations à l'échelle mondiale. L'échelle de temps universelle coordonnée (en abrégé UTC pour Universal Time Coordinated) a été créée dans ce but d'uniformisation des observations. Elle est anciennement connue sous le nom de GMT pour Greenwich Mean Time.

Les grandeurs disponibles sur le site de Météo-France sont des grandeurs mensuelles. Voici les grandeurs considérées pour la comparaison et les condition d'obtention par Météo France :

- La moyenne mensuelle de la température maximale quotidienne  $T_x$  en  $^{\circ}C$ . Au préalable, la température la plus élevée pour chaque jour  $J$  a été relevée sur le site entre le jour  $J$  à 6 h UTC et le jour  $J + 1$  à 6 h UTC. La moyenne sur le mois est ensuite faite.
- La moyenne mensuelle de la température minimale quotidienne  $T_n$  en  $^{\circ}C$ . Au préalable, la température la plus basse pour chaque jour  $J$  a été relevée sur le site entre le jour  $J - 1$  à 18 h UTC et le jour  $J$  à 18 h UTC. La moyenne sur le mois est ensuite faite.
- La durée d'ensoleillement mensuel  $Ens$  en heure [h]. Il s'agit du décompte des heures où le rayonnement normal direct est supérieur ou égal à  $120 W.m^{-2}$

A partir de  $T_n$  et  $T_x$  la moyenne mensuelle de température  $T_m$  peut être calculée telle que (source OMM) :

$$T_m = \frac{T_n + T_x}{2} \quad (3.6)$$

**Table 3.2** – Données des normales mensuelles Météo-France en janvier et juillet de 2004 à 2013 pour la station Lyon St Exupéry (source Météo-France [40]).

Lyon	Janvier				Juillet			
	Année	$T_n$ [°C]	$T_x$ [°C]	$T_m$ [°C]	Ens [h]	$T_n$ [°C]	$T_x$ [°C]	$T_m$ [°C]
2004	1,3	6,6	4	60	16,4	27,7	22,1	286
2005	-0,1	6,7	3,3	118	17,7	28,3	23	311
2006	-0,8	5	2,1	94	20,2	32,3	26,3	330
2007	3,7	9,3	6,5	55	16,3	25,6	21,0	243
2008	2,5	9,6	6,1	99	15,9	26,8	21,4	296
2009	-1,7	4,7	1,5	87	17,5	28,5	23	288
2010	-1,2	3,3	1,1	52	18,1	29,6	23,9	324
2011	0,9	7	4	86	14,8	24,5	19,7	231
2012	2,4	7,2	4,8	43	15,7	26,6	21,2	296
2013	0,7	6,1	3,4	42	18,1	29,7	23,9	322
<b>Moyenne</b>	0,77	6,55	3,68	73,6	17,1	28,0	22,5	293

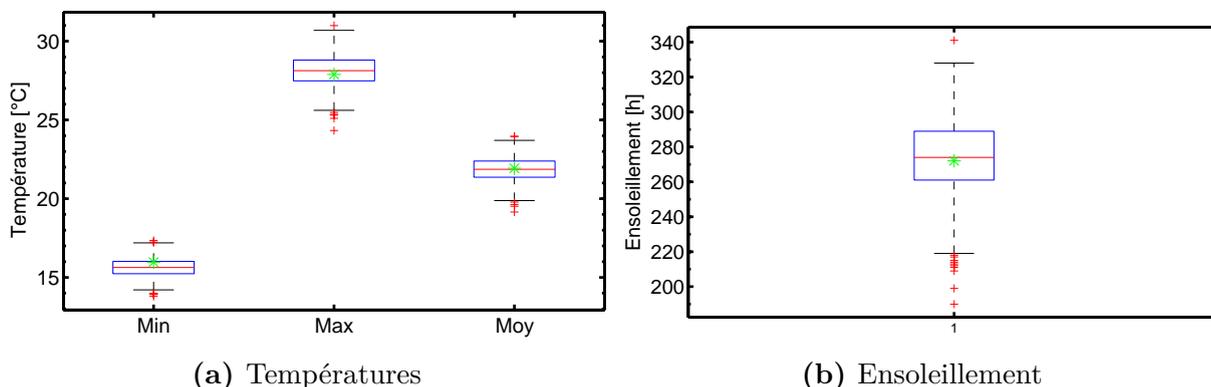
Une période de dix ans, de 2004 à 2013, de ces grandeurs mensuelles est disponible par Météo-France [40].

### 3.3.4.1 Extraction des normales mensuelles des échantillons générés

On extrait des  $N$  échantillons générés les grandeurs correspondantes aux observations disponibles. Pour chaque mois artificiel généré, les normales sont calculées. On note :

- $\mathbf{T}_n^{\text{tech}}$  les  $N$  moyennes mensuelles de la température minimale quotidienne.
- $\mathbf{T}_x^{\text{tech}}$  les  $N$  moyennes mensuelles de la température maximale quotidienne.
- $\mathbf{T}_m^{\text{tech}}$  les  $N$  moyennes mensuelles.
- $\mathbf{Ens}^{\text{ech}}$  les  $N$  valeurs d'ensoleillement mensuelles.

De la même manière on extrait du fichier IWEC source les mêmes informations telles que :  $T_n^{\text{iwec}}, T_x^{\text{iwec}}, T_m^{\text{iwec}}, Ens^{\text{iwec}}$ . Sachant que le temps légal en France est à un temps UTC+1, on prend garde de calculer  $T_n$  et  $T_x$  respectivement, du jour  $J$  7 h au jour  $J + 1$  7 h, et du jour  $J - 1$  19 h au jour  $J$  19 h. Les valeurs sont regroupées pour chaque grandeur sous forme de boxplot pour visualiser l'amplitude de la dispersion ainsi que sa symétrie éventuelle. La valeur de référence (signal source) est représentée en étoile verte. Les périodes de juillet et janvier sont étudiées pour Lyon.

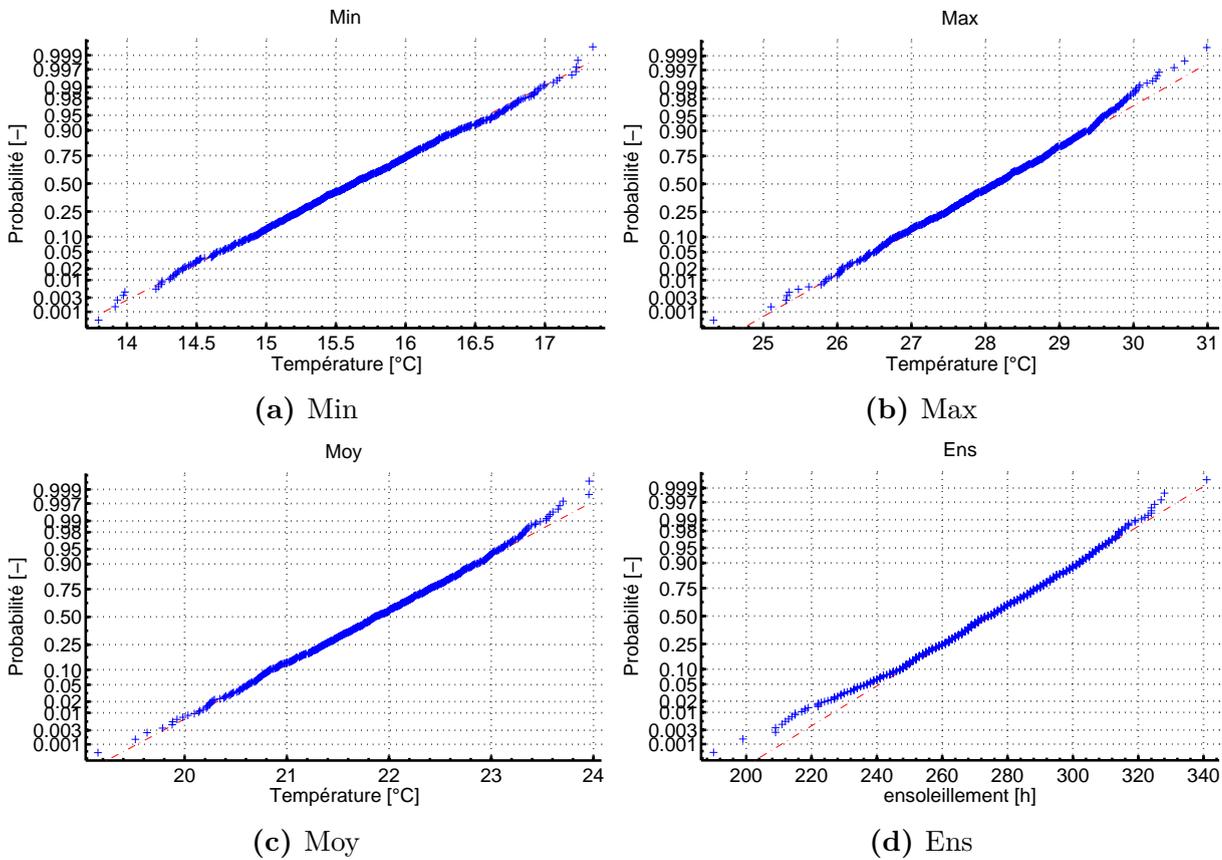


**Figure 3.11 – Lyon en juillet.** Dispersion des 1000 valeurs pour les 4 normales mensuelles sous forme de boxplot. L'étoile verte représente la valeur calculée pour le signal source pour chaque normale.

- a) Température minimale quotidienne moyenne  $T_n$ , le maximum journalier moyen  $T_x$  et la moyenne journalière  $T_m$
- b) Ensoleillement  $Ens$

### 3.3.4.2 Résultats pour Lyon au mois de juillet

La figure 3.11 regroupe les normales mensuelles pour les échantillons générés pour le climat Lyon en juillet. Les dispersions sont symétriques et la valeur de référence est bien en accord avec la médiane. Pour caractériser la dispersion, la droite de Henri est tracée pour chaque grandeur en figure 3.12. Cette représentation permet de visualiser l'écart à la loi normale de la dispersion. C'est également une façon rapide d'évaluer les valeurs en fonction des différents niveaux de probabilité. Par exemple pour la figure 3.12a on estime le centre de la distribution autour de  $15,5^{\circ}\text{C}$  et la probabilité d'obtenir des valeurs entre approximativement  $14,3^{\circ}\text{C}$  et  $17^{\circ}\text{C}$  est d'environ 98%. On observe sur la figure 3.12d un écart à la normale sur les valeurs basses qui tendent vers des probabilités plus importantes que celles de queue de gaussienne, cela signifie que les valeurs de bas ensoleillement sont plus probables que celles d'ensoleillement élevées. De même, selon la figure 3.12b, il est plus probable d'observer des valeurs hautes que basses pour la température maximale journalière moyenne. Afin de caractériser les échantillons, nous faisons l'hypothèse d'une distribution gaussienne et estimons l'écart type et la moyenne de chaque dispersion. Les résultats sont dans le tableau 3.3.



**Figure 3.12 – Lyon en juillet** Écart à la loi normale pour a) la température minimale quotidienne moyenne ( $T_n$ ), b) la température maximale quotidienne moyenne ( $T_x$ ), c) la température moyenne quotidienne moyenne ( $T_m$ ) et d) l’ensoleillement quotidien moyen  $Ens$ .

**Table 3.3 – Lyon en juillet.** Moyenne, Écart type, des échantillons et intervalle de confiance au seuil de probabilité de 95%. Les écarts relatifs à la moyenne pour 95% pour les températures sont de l’ordre de 7% (1,2°C, 1,9°C, 1,5°C) et de 16% (43 h) pour l’ensoleillement.

Grandeur	Moyenne $\mu$	Écart type $\sigma$	Intervalle de confiance pour 95%
$T_n$ [°C]	15,6	0,589	$15,6 \pm 1,2$
$T_x$ [°C]	28,1	0,969	$28,1 \pm 1,9$
$T_m$ [°C]	21,9	0,755	$21,9 \pm 1,5$
$Ens$ [°C]	274	21,8	$274 \pm 43$

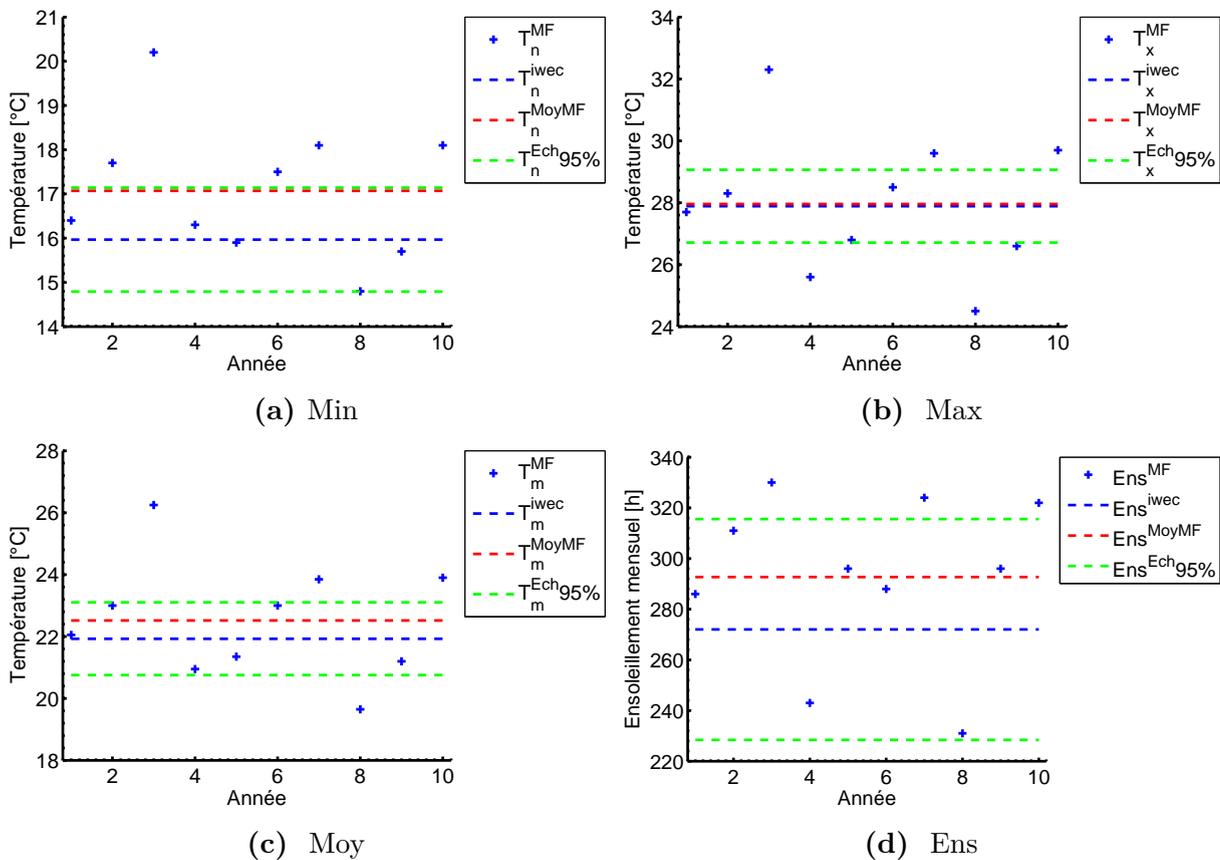
On observe dans le tableau 3.3 un écart relatif à la moyenne pour 95% de niveau de confiance, de 7% pour la température moyenne ( $21,9 \pm 1,5$ ) et de 16% ( $274 \pm 43$ ) pour l’ensoleillement. Les valeurs de dispersion obtenues sont comparées aux valeurs relevées

par Météo-France. Les données dont nous disposons sont limitées, le but de cette comparaison est de vérifier si l'ordre de grandeur de la dispersion que l'on a générée est cohérent avec la variabilité météo. En figure 3.13 La comparaison est faite pour les quatre normales mensuelles. Pour chaque graphe, les valeurs des mesures de 2004 à 2013 sont représentées ainsi que la moyenne de ces valeurs ( $X^{MoyMF}$ ). La moyenne Météo-France est comparée à la moyenne du fichier source ( $X^{iwec}$ ). Puis on trace l'intervalle de confiance à 95% autour de la valeur moyenne du IWEC afin de visualiser l'étendu de la dispersion des échantillons générés par rapport aux observations.

On remarque tout d'abord des écarts entre la valeur moyenne de MF et du IWEC, particulièrement pour  $T_n$  et  $Ens$ . Ces écarts peuvent être dus à la différence de période de prise de mesure car le IWEC est un fichier construit à base d'observations de 1982 à 1999. Comme la période d'observations de MF est plus courte, il est possible que des valeurs extrêmes hautes "tirent" la moyenne de MF vers le haut comme nous l'observons. Par exemple les valeurs de températures mensuelles observées pour l'année 2006 (3ième point) sont plus élevées que le reste des observations. En ce qui concerne l'ensoleillement la valeur de 2006 n'est pas particulièrement détachée du reste des observations. D'autre part, l'ensoleillement semble sous évalué dans les échantillons par rapport aux mesures.

Malgré ces écarts de moyenne, la dispersion à 95% ( $X^{Ech95\%}$ ) des échantillons englobe au moins la moitié des observations. Cela paraît en accord avec le fait que les échantillons sont générés à partir d'un fichier moyen. Les dispersions sont donc cohérentes avec la réalité et le caractère moyen du IWEC. Bien que plus étendue, la dispersion liée à l'ensoleillement reste dans l'ordre de grandeur de la variabilité naturelle avec sept observations sur les dix qui sont comprises dans l'intervalle.

Nous utilisons le même protocole de comparaison pour le mois de janvier.



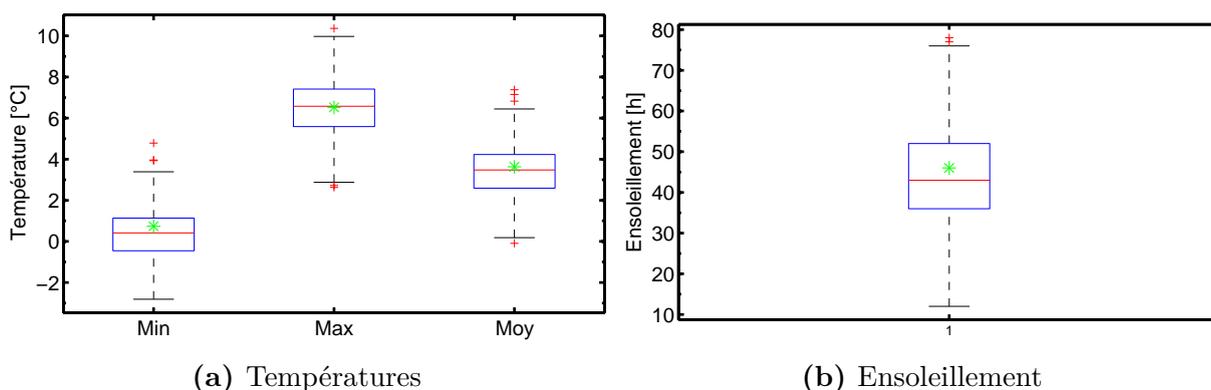
**Figure 3.13 – Lyon en juillet** Comparaison des observations de Meteo-France par rapport à la dispersion des échantillons. Chaque graphe est composé de 10 ans d’observations MF de 2004 à 2013 (croix bleues), de la moyenne MF (ligne pointillée rouge) ainsi que la moyenne du IWEC (ligne pointillée bleue). La dispersion des échantillons est représentée par l’intervalle à  $1.96\sigma$  (cf. valeur tableau 3.3) autour de la moyenne IWEC (lignes pointillées vertes).

- a) Température minimale quotidienne moyenne ( $T_n$ )
- b) Température maximale quotidienne moyenne ( $T_x$ )
- c) Température moyenne quotidienne moyenne ( $T_m$ )
- d) Ensoleillement quotidien moyen ( $Ens$ )

### 3.3.4.3 Résultats pour Lyon au mois de janvier

Nous appliquons la même démarche que pour le cas de Lyon en juillet. Tout d’abord les dispersions sont évaluées en figure 3.14, la valeur de référence est en accord avec la médiane de chaque dispersion. Les caractéristiques statistiques pour chaque dispersion sont répertoriées dans le tableau 3.4. On observe une dispersion plus importante en janvier qu’en juillet (cf. tableau 3.3). Cette différence d’amplitude est cohérente avec les phénomènes météorologiques, en effet, il existe une saisonnalité dans la variabilité des

températures [53] et la variabilité naturelle en hiver est plus importante qu'en été dans les climats tempérés.

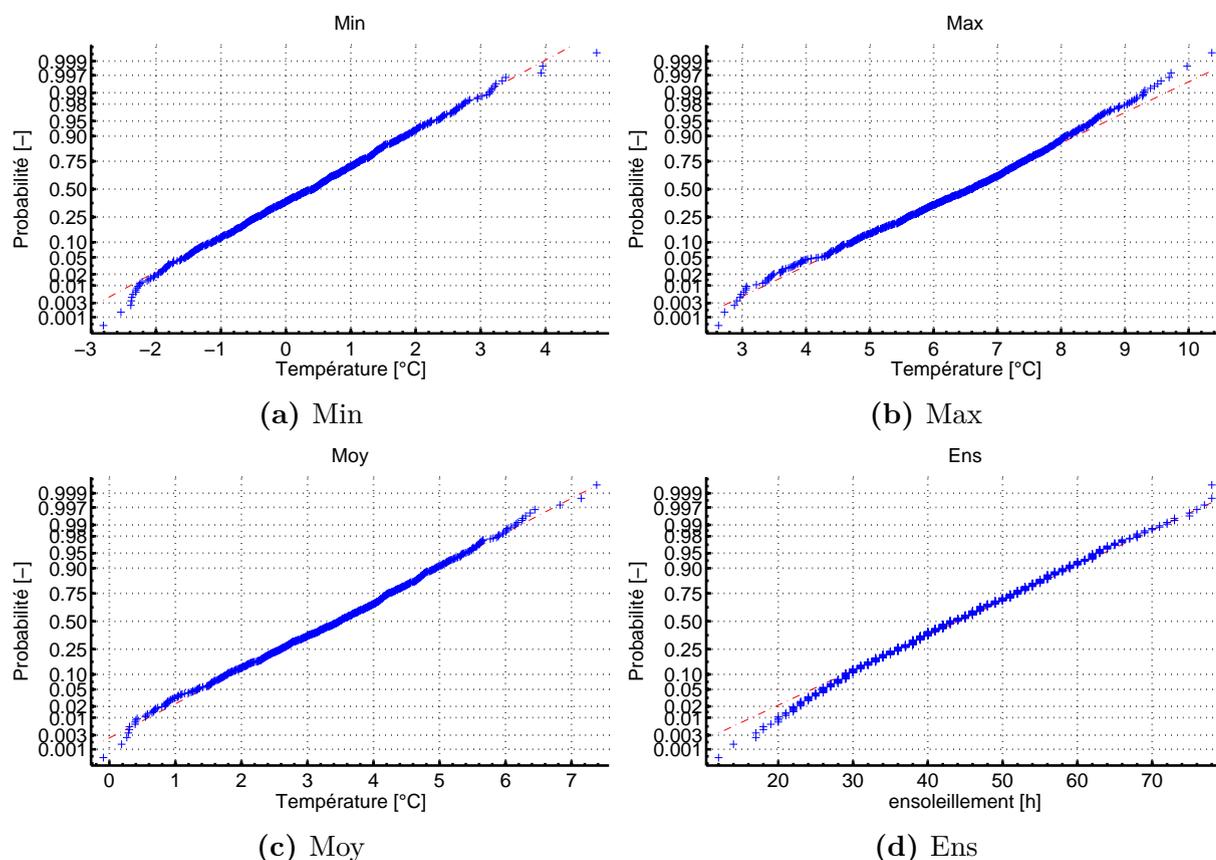


**Figure 3.14 – Lyon en janvier** Dispersion des 1000 valeurs pour les 4 normales mensuelles par boxplot avec en a) la température minimale quotidienne moyenne  $T_n$ , le maximum journalier moyen  $T_x$  et la moyenne journalière  $T_m$  et en b) l'ensoleillement  $Ens$ . L'étoile verte représente la valeur calculée pour le signal source pour chaque normale.

**Table 3.4 – Lyon en janvier** Écart type, moyenne des échantillons et intervalle de confiance au seuil de probabilité de 95%. La variabilité est plus importante qu'en juillet (cf. tableau 3.3) la valeur de l'écart type est presque doublée

Grandeur	Moyenne $\mu$	Écart type $\sigma$	Intervalle de confiance pour 95%
$T_n$ [°C]	0,360	1,16	$0,36 \pm 2,27$
$T_x$ [°C]	6,48	1,32	$6,48 \pm 2,59$
$T_m$ [°C]	3,42	1,23	$3,42 \pm 2,41$
$Ens$ [°C]	43,9	11,2	$43,9 \pm 22,0$

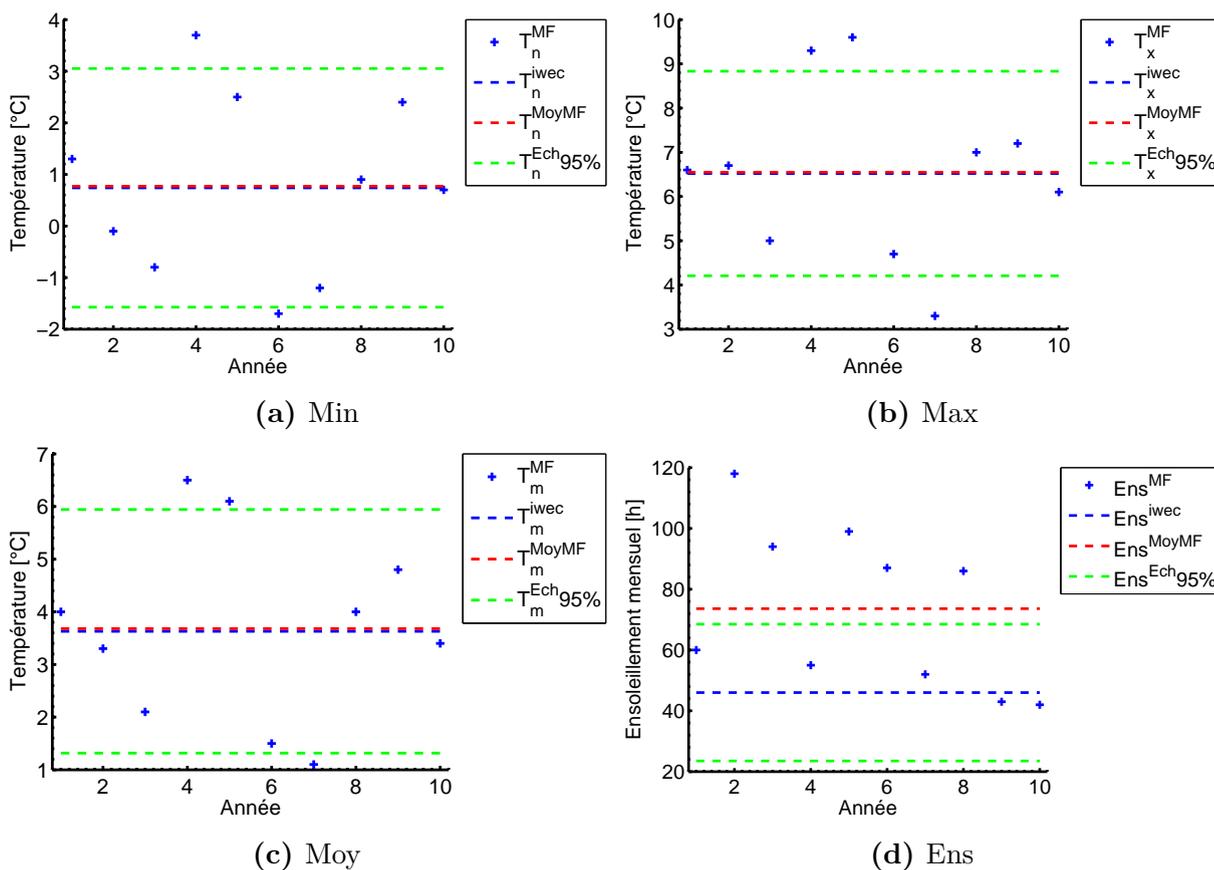
Le tableau 3.4 montre une variabilité pour les températures plus importante en hiver qu'en été (cf. tableau 3.3). Par exemple la température moyenne a un écart type de 1,23 en hiver contre 0,75 en été.



**Figure 3.15 – Lyon en janvier.** Écart à la loi normale pour a) la température minimale quotidienne moyenne ( $T_n$ ), b) la température maximale quotidienne moyenne ( $T_x$ ), c) la température moyenne quotidienne moyenne ( $T_m$ ) et d) l’ensoleillement quotidien moyen  $Ens$ .

Nous vérifions en figure 3.16 l’étendue de la dispersion par rapport aux observations Météo-France de 2004 à 2013. Tout d’abord, nous observons que contrairement à l’étude sur le mois de juillet les valeurs des moyennes des observations MF et IWEC pour les températures coïncident. En ce qui concerne l’ensoleillement, celui-ci est sous évalué comme en juillet. Le décalage vers le bas entre les moyennes MF et IWEC est du même ordre de grandeur en juillet et janvier, soit d’environ 20 h d’ensoleillement. Par ailleurs, pour les températures, 95% de la dispersion des échantillons englobe au moins sept années sur les dix ans d’observations, la variabilité des échantillons en hiver est donc plus importante que celle observée en été. Pour l’ensoleillement, les conclusions sont similaires à celles de juillet, puisque l’ensoleillement est sous évalué dans le IWEC par rapport aux observations de 2004 à 2013, la dispersion est en décalage avec les observations, elle englobe tout de même 50% des observations et l’intervalle de confiance de la dispersion est du même ordre de grandeur que la variabilité des observations d’ensoleillement. Les échantillons générés

sont représentatifs de la variabilité des observations.



**Figure 3.16 – Lyon en janvier** Comparaison des observations de Meteo-France par rapport à la dispersion des échantillons. Chaque graphe est composé de dix ans d’observations MF de 2004 à 2013 (croix bleues), de la moyenne MF (ligne pointillée rouge) ainsi que la moyenne du IWEC (ligne pointillée bleue). La dispersion des échantillons est représentée par l’intervalle à  $1.96\sigma$  (cf. valeur tableau 3.3) autour de la moyenne IWEC (lignes pointillées vertes).

- a) Température minimale quotidienne moyenne ( $T_n$ )
- b) Température maximale quotidienne moyenne ( $T_x$ )
- c) Température moyenne quotidienne moyenne ( $T_m$ )
- d) Ensoleillement quotidien moyen ( $Ens$ )

Nous avons vu que, pour Lyon sur les deux périodes, la dispersion des échantillons pour les quatre grandeurs de normales mensuelles est cohérente avec la variabilité naturelle des observations de Météo-France. Nous observons toutefois une sous-évaluation des valeurs de l’échantillon d’ensoleillement. Afin de vérifier si cette sous évaluation est caractéristique de Lyon, une autre localisation est étudiée ce qui permettra également de montrer l’aspect générique de la méthode présentée. Pour cela le climat de Strasbourg est étudié.

### 3.4 Application à une autre localisation : la ville de Strasbourg

#### 3.4.1 Normales mensuelles de Météo-France

Dans le tableau 3.5 les normales mensuelles de Météo-France de 2004 à 2013 pour Strasbourg sont répertoriées.

**Table 3.5** – Données des normales mensuelles Météo-France de 2004 à 2013 pour la ville de Strasbourg. (source Météo-France [40])

Strasbourg	Janvier				Juillet			
	$T_n$ [°C]	$T_x$ [°C]	$T_m$ [°C]	$Ens$ [h]	$T_n$ [°C]	$T_x$ [°C]	$T_m$ [°C]	$Ens$ [h]
2004	-0,2	5,1	2,5	46	14,1	25,3	19,7	211
2005	0,2	6,2	3,2	76	15	25,6	20,3	203
2006	-3,2	1,6	-0,8	72	17,4	31,1	24,3	331
2007	3,3	8,8	6,1	36	14,4	23,9	19,2	171
2008	1	8	4,5	71	14,4	25,5	20	211
2009	-4,6	2	-1,3	103	14,8	25,5	20,2	214
2010	-2,6	1	-0,8	24	15,7	27,8	21,8	272
2011	-0,1	5,5	2,7	59	12,4	23,2	17,8	205
2012	1,4	6,8	4,1	58	14,4	24,9	19,7	208
2013	0,1	4,3	2,2	19	15,9	28,3	22,1	304
<b>Moyenne</b>	-0,47	4,93	2,23	56,4	14,9	26,1	20,5	233

#### 3.4.2 Comparaison des échantillons

Des échantillons sont générés sur les deux périodes, soit deux séries de 1000 échantillons. Pour chaque jeu d'échantillons le tableau regroupant les caractéristiques statistiques est présenté puis les graphes permettant la comparaison avec les observations de Météo-France sont commentés.

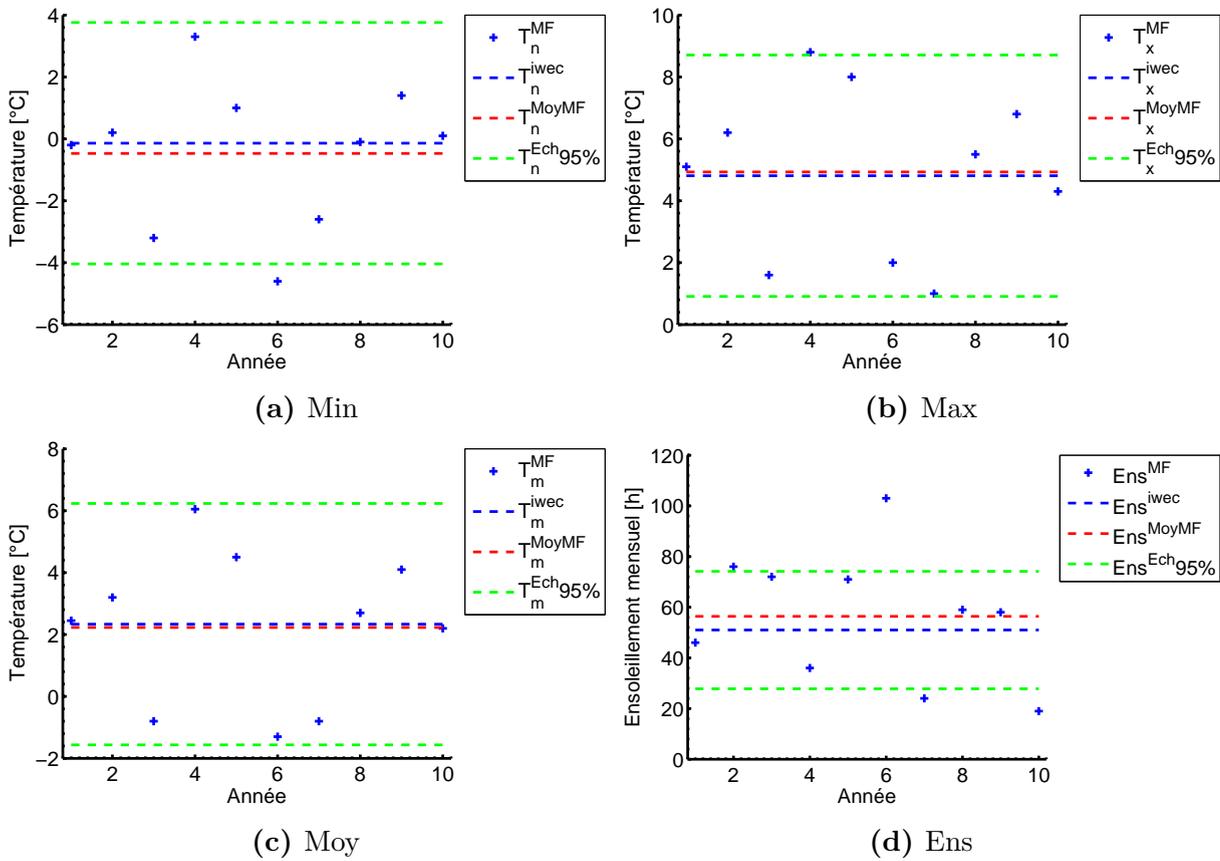
##### 3.4.2.1 Résultats pour Strasbourg au mois de janvier

Les dispersions pour chaque grandeur sont renseignées dans le tableau 3.6. On observe des températures inférieures à celles de Lyon sur la même période et une dispersion plus importante (cf. tableau 3.4). Cette dispersion importante se visualise sur la figure 3.17 où les observations Météo-France et les échantillons sont comparés. Pour les normales de température, presque la totalité des observations est comprise dans l'intervalle de confiance

de 95%. Tout comme la même période de janvier pour Lyon, les moyennes Météo-France et IWEC sont proches. Contrairement à Lyon, on observe également des valeurs proches pour l'ensoleillement. La variabilité de l'ensoleillement est plus réduite que celle de la température, dans le sens où seule un peu plus de la moitié des observations est comprise dans l'intervalle de confiance. La dispersion des échantillons de température est représentative de la variabilité complète des observations sur dix ans.

**Table 3.6 – Strasbourg en janvier** Moyenne et écart type des échantillons et intervalle de confiance au seuil de probabilité de 95%.

Grandeur	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance pour 95%
$T_n$ [ $^{\circ}C$ ]	-0,02	1,95	$-0,02 \pm 3,82$
$T_x$ [ $^{\circ}C$ ]	4,62	1,99	$4,62 \pm 3,90$
$T_m$ [ $^{\circ}C$ ]	2,30	1,97	$2,30 \pm 3,86$
$Ens$ [h]	49,2	11,6	$49,2 \pm 22,7$



**Figure 3.17 – Strasbourg en janvier.** Comparaison des observations de Météo-France par rapport à la dispersion des échantillons. Chaque graphe est composé de dix ans d’observations MF de 2004 à 2013 (croix bleues), de la moyenne MF (ligne pointillée rouge) ainsi que la moyenne du IWEC (ligne pointillée bleue). La dispersion des échantillons est représentée par l’intervalle à  $1.96\sigma$  (cf. valeur tableau 3.6) autour de la moyenne IWEC (lignes en pointillés verts).

- a) Température minimale quotidienne moyenne ( $T_n$ )
- b) Température maximale quotidienne moyenne ( $T_x$ )
- c) Température moyenne quotidienne moyenne ( $T_m$ )
- d) Ensoleillement quotidien moyen ( $Ens$ )

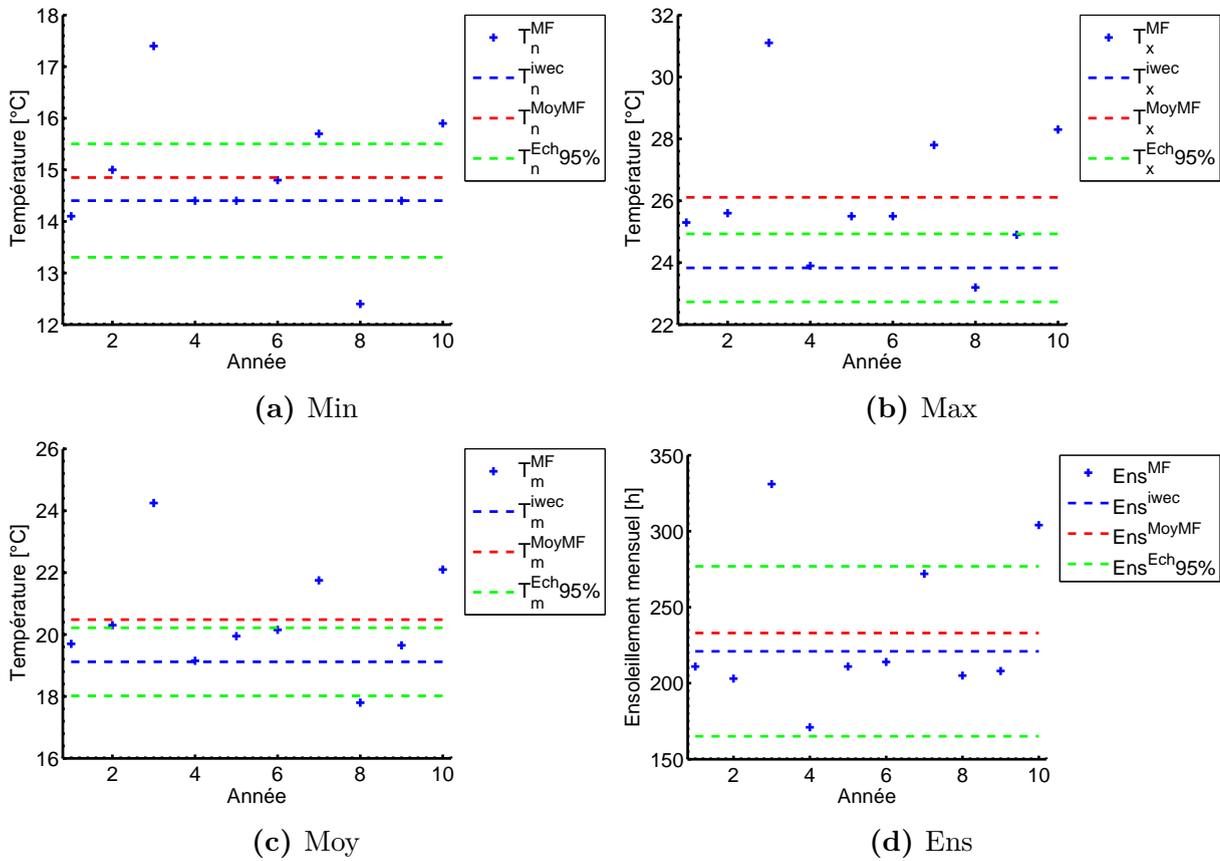
### 3.4.2.2 Résultats pour Strasbourg au mois de juillet

Le tableau 3.7 regroupe les valeurs mensuelles pour le mois de juillet à Strasbourg. On observe des dispersions moins importantes qu’en hiver et du même ordre de grandeur que celles de Lyon sur la même période (tableau 3.3).

**Table 3.7 – Strasbourg en juillet** Moyenne, écart type des échantillons et intervalle de confiance au seuil de probabilité de 95%.

<b>Grandeur</b>	<b>Moyenne <math>\mu</math></b>	<b>Ecart type <math>\sigma</math></b>	<b>Intervalle de confiance pour 95%</b>
$T_n$ [ $^{\circ}C$ ]	14,4	0,55	$14,4 \pm 1,1$
$T_x$ [ $^{\circ}C$ ]	23,6	0,68	$23,6 \pm 1,3$
$T_m$ [ $^{\circ}C$ ]	19,0	0,59	$19,0 \pm 1,2$
$Ens$ [h]	222	28,0	$222 \pm 55$

La comparaison avec les normales de Météo-France est représentée en figure 3.18. De la même manière que pour Lyon à la même période (figure 3.13), on observe un décalage entre la moyenne Météo-France et la moyenne IWECC. Le fichier moyen sous-estime la valeur moyenne MF par rapport aux observations et particulièrement pour la valeur de température maximale quotidienne moyenne. Dans le cas de Strasbourg le décalage pour l'ensoleillement est moins important que pour Lyon. L'ordre de grandeur de l'intervalle de confiance est cohérent avec la dispersion des observations. Mis à part le cas de la température maximale où l'échantillon sous-estime les valeurs, la dispersion couvre plus de la moitié des observations. On dispose, encore une fois, d'un échantillon d'une variabilité cohérente avec la variabilité naturelle moyenne.



**Figure 3.18 – Strasbourg en juillet.** Comparaison des observations de Météo-France par rapport à la dispersion des échantillons. Chaque graphe est composé de dix ans d’observations MF de 2004 à 2013 (croix bleues), de la moyenne MF (ligne pointillée rouge) ainsi que la moyenne du IWEC (ligne pointillée bleue). La dispersion des échantillons est représentée par l’intervalle à  $1.96\sigma$  (cf. valeur tableau 3.3) autour de la moyenne IWEC (lignes pointillées vertes).

- a) Température minimale quotidienne moyenne ( $T_n$ )
- b) Température maximale quotidienne moyenne ( $T_x$ )
- c) Température moyenne quotidienne moyenne ( $T_m$ )
- d) Ensoleillement quotidien moyen  $Ens$

### 3.5 Des échantillons cohérents : vers l’analyse de sensibilité

Les exemples de génération présentés ici font partie d’une démarche exploratoire sur la gestion des analyses de sensibilité et d’incertitude lorsque les entrées incertaines sont auto-corrélées et inter-corrélées. Nous avons voulu générer une dispersion cohérente avec la variabilité naturelle pour chaque variable et respecter des contraintes d’indépendance

de génération aléatoire. Le but de la démarche est de vérifier que les échantillons que l'on génère ne sont pas aberrants et respectent la physique sous-jacente des données météorologiques ainsi que la variabilité naturelle.

Nous avons montré que les corrélations entre chaque variable échantillonnée reste dans l'ordre de grandeur et dans le sens physique par rapport aux corrélations du signal source. Ces corrélations sont toutefois perturbées dans les échantillons car il y a une perte d'informations par la méthode d'échantillonnage ce qui peut induire des incohérences dans la génération et notamment pour les rayonnements direct et diffus qui sont conditionnés par la nébulosité.

Nous avons ensuite comparé les échantillons à des observations Météo-France de normales mensuelles pour la température et l'ensoleillement. Certaines grandeurs peuvent être sous-évaluées mais l'ordre de grandeur de la dispersion engendrée par l'échantillonnage est cohérent avec la variabilité naturelle de la météo. En fonction des localisations et de la période cette variabilité est plus ou moins importante, par exemple la variabilité de la température en hiver est plus importante pour Strasbourg que pour Lyon.

Pour des raisons de manque de données mesurées sur le long terme la comparaison est restée au niveau de l'ordre de grandeur. Pour qualifier la représentativité de la variabilité météo pour un climat donnée, il serait intéressant de comparer les échantillons à des observations horaires sur le long terme.

Nous avons voulu montrer les possibilités de la méthodologie développée qui peut être appliquée à différents types d'entrées fonctionnelles en extrayant la fonction de distribution, la moyenne et la fonction d'auto-corrélation du signal. Dans le cadre de ce travail la dispersion que nous avons souhaitée représenter est celle de la variabilité temporelle sur le long terme des données météorologiques. La difficulté a été en premier lieu de caractériser cette variabilité pour pouvoir l'appliquer mais on peut imaginer utiliser la méthodologie développée pour imposer la dispersion que l'on souhaite en choisissant la fonction de distribution comme par exemple des problématiques d'incertitude de mesure sur des entrées fonctionnelles.



# Chapitre 4

## Estimation de l'incertitude des besoins énergétiques d'un bâtiment liée à la variabilité météorologique

---

Dans ce chapitre, on utilise les échantillons des données météorologiques pour réaliser l'analyse d'incertitude et de sensibilité. On vérifie tout d'abord la cohérence des résultats à la fois des dispersions et de la sensibilité et ce pour les deux climats de Strasbourg et Lyon. Puis on teste les incertitudes de sortie et la sensibilité sur des variantes de modèle thermique de la maisons INCAS. Enfin on réalise une analyse de sensibilité et d'incertitude en prenant en compte entrées statiques et dynamiques.

---

### 4.1 Protocole

#### 4.1.1 Objectif et contexte de l'étude

A partir de la méthode d'échantillonnage du chapitre 3, nous avons généré des échantillons pour chaque variable météorologique. Chaque ensemble est constitué de six matrices correspondant aux six variables météo considérées, et chaque matrice comporte 1000 tirs (simulations) de 744 valeurs. Ces 1000 évaluations représentent la variabilité météorologique du climat source.

Nous souhaitons obtenir à partir de ces échantillons la réponse d'un modèle de bâtiment par simulation thermique dynamique. Ce modèle est une maison basse consommation de stratégie d'économie d'énergie dite "passive" (cf. section cas test 1.2.1) dont l'isolation

des murs, des planchers et des plafonds ainsi que l'inertie est importante. Le modèle est simplifié notamment au niveau du système qui est pris comme idéal et aucun scénario n'est considéré, il n'y a pas de renouvellement d'air et seules des faibles infiltrations sont prises en compte. Nous souhaitons en effet dans cette étude, vérifier la réponse du bâtiment au niveau de l'enveloppe et minimiser les incertitudes liées au système complexe réel de cette maison qui est une VMC double flux. Il s'agit au niveau des réponses de ce bâtiment d'étudier les besoins de chauffage et de refroidissement, qui ne sont pas représentatifs de la consommation réelle du bâtiment en fonctionnement. On souhaite par cette étude montrer sur un modèle simple la méthodologie de propagation des incertitudes et d'estimation de la sensibilité (de la variable météorologiques responsable de cette incertitude). L'étude de la réponse du bâtiment est une vérification de la validité de la méthode.

Pour pouvoir vérifier indirectement la méthodologie développée en analyse de sensibilité, on applique le même protocole d'estimation d'incertitude et des variables responsables de cette incertitude à deux localisations : Lyon et Strasbourg sur les deux périodes estivale et hivernale. Puis nous testons la modification technologique de composants : en remplaçant les menuiseries par du triple vitrage puis en ayant recours à un modèle où la performance du bâtiment est dégradée en diminuant l'isolation, la masse du bâtiment et en augmentant les infiltrations.

Enfin le modèle thermique de base est étudié en prenant en compte les incertitudes associées aux paramètres statiques en plus des incertitudes des données météorologiques.

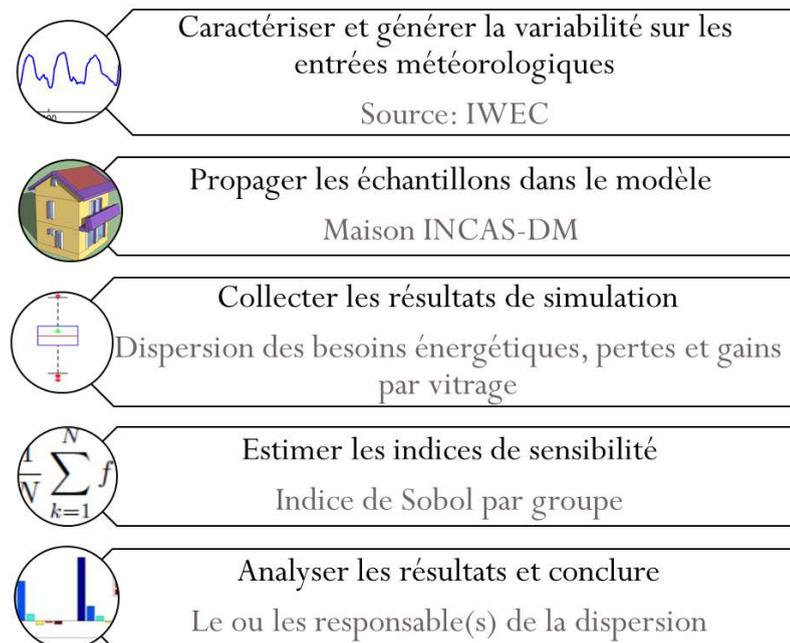
#### 4.1.2 Description de la méthode

Pour un échantillon donné de  $N \times p \times N_t$  donc dans notre cas de six matrices de 1000 par 744, on construit par permutation un second échantillon pour permettre l'analyse de sensibilité. Cette étape est présentée en section 2.2.3.1 et permet un gain en coût de calcul important en rendant le nombre de simulation indépendant du nombre de variables étudiées. Le modèle est ensuite évalué avec ces 2 échantillons, soit 2000 simulations d'où la nécessité d'automatiser les simulations pour cette étape. Dans le cadre de ce travail de thèse, une routine Matlab permet la création de fichier de type ".bat" qui code l'appel de chaque simulation par EnergyPlus avec le jeu d'échantillons correspondant et les sorties sont extraites automatiquement. Ce programme a été développé par Rabouille et al. [52] dans le cadre d'une problématique de réduction de modèle par polynôme de chaos. Le temps d'une simulation est compris entre trente secondes et une minute pour un mois simulé i.e. de 17 h à 34 h de calculs pour un processeur. Un ordinateur Quad-Core a été utilisé pour ce travail ce qui permet de réduire le temps des simulations de 4 h pour une simulation de trente secondes à 8 h pour une simulation allant jusqu'à une minute. Le

nombre important de simulations est nécessaire pour permettre la convergence des indices de sensibilité.

La dispersion sur chaque sortie nous permet d'évaluer la distribution et donc l'incertitude et nous utilisons la représentation boxplot pour la représenter, si cette dispersion est assimilable à une loi normale nous estimons moyenne et écart type ainsi que l'intervalle d'incertitude sur la performance du bâtiment à 95% de seuil de confiance. Cette démarche va permettre la comparaison de différentes configurations. Les indices de sensibilité (effet seul de chaque variable) sont ensuite estimés selon la méthode présentée en section 2.2.2.2 ce qui permettra de conclure sur la ou les variable(s) météo responsable(s) de l'incertitude de sortie. Pour chaque indice de sensibilité, on détermine la stabilité de l'estimateur par un processus de ré-échantillonnage de type bootstrap (section 2.2.3.2) ce qui consiste à ré-évaluer les indices de sensibilité pour 100 répliques bootstrap et d'en extraire la moyenne et l'écart-type.

Un schéma récapitulatif de la démarche globale est présenté en figure 4.1.



**Figure 4.1** – Récapitulatif des étapes pour l'évaluation de la sensibilité et de l'incertitude d'un modèle de bâtiment aux sollicitations météorologiques

### 4.1.3 Analyse des sorties de simulation

Pour évaluer le comportement et la performance d'un bâtiment on se base sur les besoins de chauffage et de refroidissement. En effet le modèle que nous utilisons n'est pas assez détaillé pour considérer une sortie comme le confort ou la consommation. Les besoins permettent d'obtenir une sortie scalaire représentative de la température extérieure et le besoin de chauffage total est l'agrégation des besoins de chaque zone. En plus des besoins totaux sur le bâtiment, on étudie les besoins de chauffage et de refroidissement du rez-de-jardin et de l'étage pour visualiser les différences de besoins entre zone, cela pour faciliter la vérification de la cohérence des indices de sensibilité obtenus. Les besoins correspondent à l'agrégation de la performance de plusieurs composants comme par exemple celle des vitrages. On étudie la dispersion et les indices de sensibilité sur ce composant simple pour vérifier la cohérence et la sensibilité du modèle à la variation en entrée sur une sortie simple ou l'interprétation de l'influence des variables est plus aisée qu'une sortie agrégée somme de nombreux composants. Ces sorties simples vont être étudiées systématiquement pour les premiers résultats afin de mettre en valeur la cohérence des résultats de dispersion et de sensibilité en différenciant les besoins et le composant fenêtre.

#### 4.1.3.1 Le besoin de chauffage et de refroidissement

On normalise les résultats de besoins totaux pour avoir une unité proche des  $kW.m^{-2}.an^{-1}$  qui sont des  $kW.m^{-2}.mois^{-1}$  qui est l'unité usuelle des performances énergétiques. On étudie ensuite séparément les deux zones du rez-de-jardin (RDJ) et de l'étage (Etage) qui ont la même surface.

#### 4.1.3.2 Analyse des pertes et des gains par vitrage

La méthodologie mise en place est exploratoire, on étudie les gains et pertes par fenêtre pour permettre de vérifier indirectement la cohérence des résultats. Lors de l'étude de sensibilité de sorties agrégées comme les besoins, il peut être intéressant de visualiser un des composants menant à ces besoins. Le composant fenêtre est d'autant plus intéressant qu'il est sensible à la fois à la température et au rayonnement direct. Pour plus de facilité et pour limiter les sorties, on a regroupé les flux pour chaque orientation et on normalise par la surface totale de surface vitrée, ce qui permet de comparer par orientation. Pour les deux zones, on étudie les pertes et gains énergétiques pour les quatre orientations soit seize sorties supplémentaires par rapport aux trois besoins de chauffage (RDJ, Etage, total).

La première comparaison concerne deux localisations différentes avec les échantillons étudiés en chapitre 3 : Strasbourg et Lyon pour les deux périodes de janvier et juillet soit quatre ensembles d'échantillons de 2000 simulations propagés. Le climat de Lyon sert

ensuite de référence pour le reste des comparaisons sur les modèles. La première étude est particulièrement détaillée, notamment pour vérifier la cohérence des résultats de dispersion et de sensibilité. Pour des soucis de clarté face aux différents résultats apportés dans la suite seuls les besoins de chauffage et de refroidissement seront étudiés.

On rappelle le but de ces simulations :

- Caractériser la dispersion en sortie à partir des échantillons cohérents de la météo
- Quantifier l'influence de chaque variable météo sur l'incertitude de sortie

## 4.2 Comparaison de Lyon et Strasbourg : impact du climat étudié

### 4.2.1 Réponse du bâtiment en période hivernale

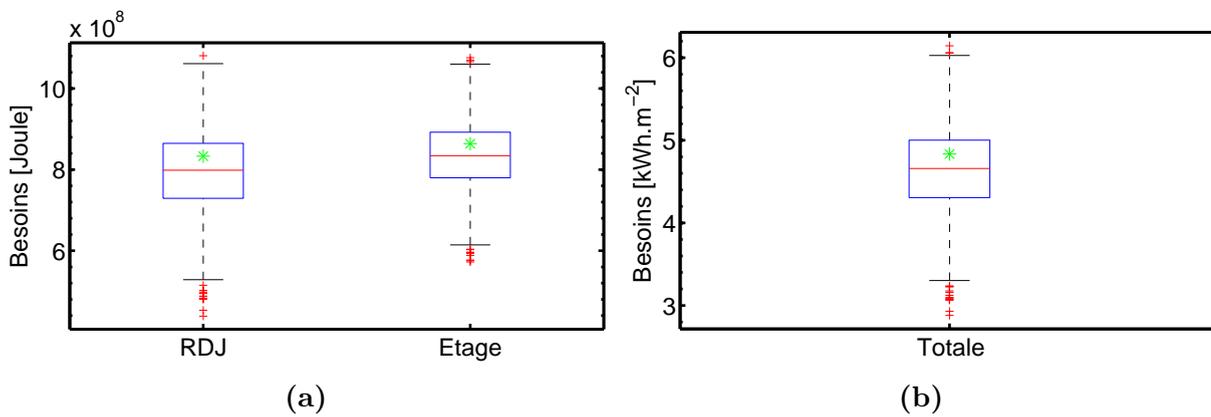
On compare les résultats de simulation de Lyon et Strasbourg pour le mois de janvier. Lyon étant la simulation de référence qui va servir de comparaison à d'autres cas de simulation. On expose tout d'abord les résultats de Lyon pour le mois de janvier, puis les résultats relatifs à Strasbourg sont comparés à ceux de Lyon.

#### 4.2.1.1 Résultats obtenus pour le climat de Lyon

##### Dispersion des résultats

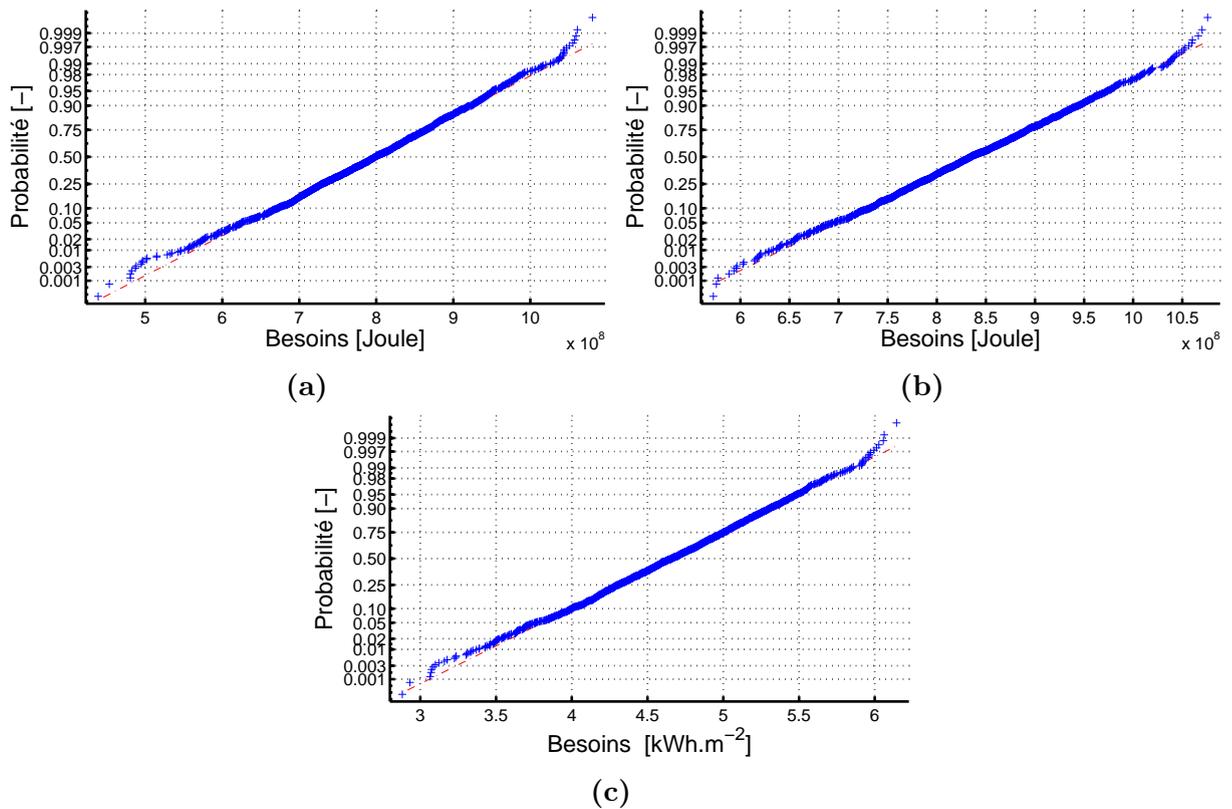
Le but est de visualiser la dispersion des 2000 évaluations du modèle. A cet effet, on va tout d'abord visualiser les boxplot, pour comparer les différentes dispersions sur les seize sorties relatives au composant fenêtre du bâtiment et pour visualiser la symétrie de la distribution, le placement de la valeur de référence par rapport à cette distribution et les valeurs extrêmes.

Tout d'abord on observe que les valeurs de référence correspondent bien à la médiane de la dispersion empirique qui est symétrique. La figure 4.3 compare la dispersion des résultats à la droite de Henri, cela permet de visualiser l'écart à la loi normale. L'hypothèse de normalité est satisfaisante, particulièrement pour des valeurs comprises entre la probabilité 0.025 et 0.975 soit 95% de l'information. Nous utilisons donc la moyenne et l'écart type pour définir les dispersions.



**Figure 4.2 – Lyon en janvier.** Dispersion des 2000 valeurs de besoins de chauffage. Représentation sous forme de boxplot avec l'étoile verte comme marque de la simulation de référence effectuée avec le IWEC source.

- a) Besoins sur les deux zones : le rez-de-jardin et l'étage
- b) Besoins de chauffage total



**Figure 4.3 – Lyon en janvier.** Écart à la normale (droite de Henri) pour les besoins de chauffage.

- a) Besoins pour le rez-de-jardin
- b) Besoins pour l'étage
- c) Besoins de chauffage total

A partir des valeurs des écarts types et des moyennes, on obtient l'intervalle de confiance correspondant à un seuil de 95% soit  $1,96\sigma$ . Pour mieux visualiser les incertitudes, la variation en % de l'écart à la valeur moyenne est aussi renseignée, en considérant les bornes associées à une probabilité de 95%. Toutes ces informations sont regroupées dans le tableau 4.1 pour le RDJ, l'Etage et le besoin total de chauffage.

La dispersion sur le besoin de chauffage due à la variabilité météo s'élève à plus de 20% (21,7% tableau 4.1) de la valeur obtenue avec le fichier IWEC. On observe entre le RDJ et l'Etage un écart de performance, avec des besoins plus élevés à l'Etage et donc un incertitude relative plus faible. Les premiers résultats montrent que l'incertitude due à la variabilité météorologique n'est pas négligeable. On rappelle que la variabilité induite en entrée est représentative à 95% d'observations réelles moyennes (5/10).

**Table 4.1 – Lyon en janvier** Valeurs de dispersion du besoin de chauffage total et par zone

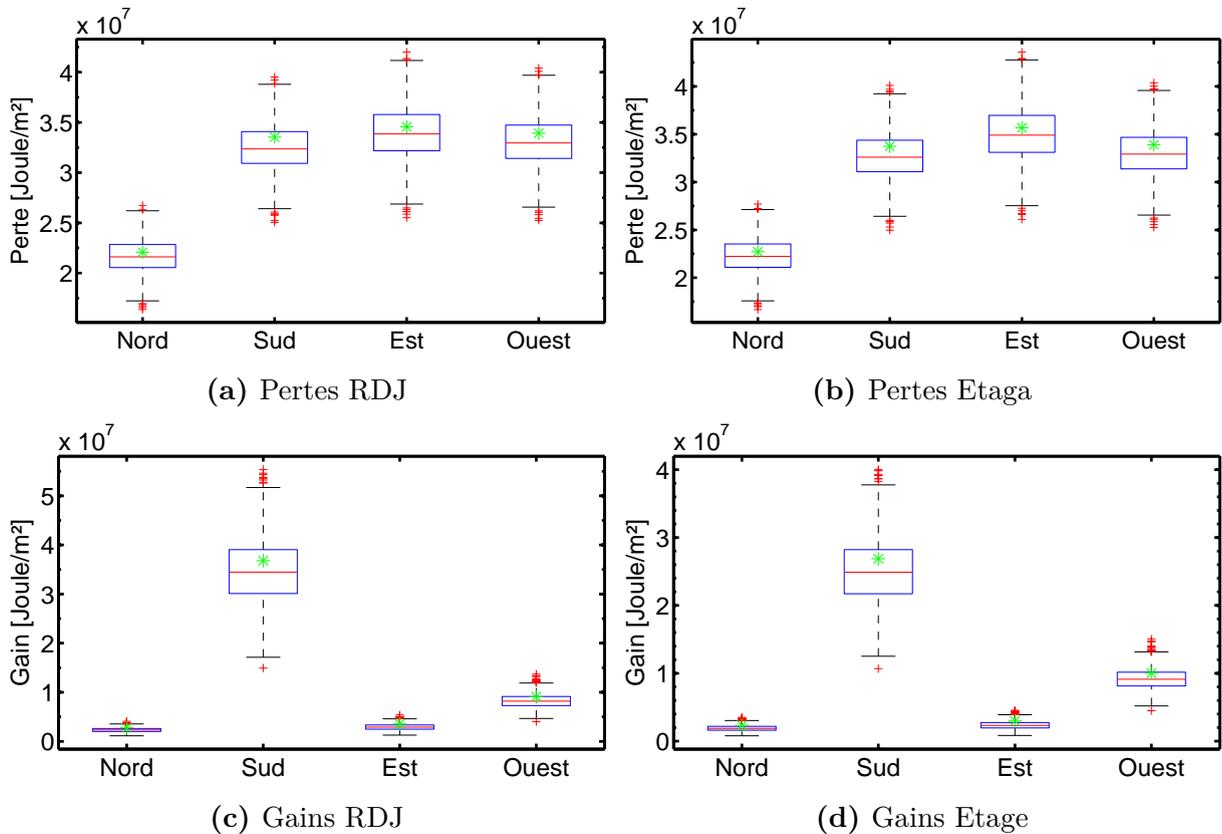
Chauffage	Moyenne $\mu$	Écart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $.10^8$ Joule]	7,96	0,990	$7,96 \pm 1,94$	24,4%
Etage [ $.10^8$ Joule]	8,36	0,841	$8,36 \pm 1,65$	19,7%
Totale [ $kWh.m^{-2}$ ]	4,65	0,516	$4,65 \pm 1,01$	21,7%

On souhaite visualiser les pertes et gains en fonction des orientations pour vérifier la cohérence des simulations et de la propagation des échantillons.

La figure 4.4 regroupe les dispersions pour chaque orientation des pertes et gains par fenêtre pour chaque zone. On rappelle que ces pertes et gains sont normalisés par la surface vitrée sur chaque orientation. L'interprétation des résultats en est plus aisée, puisque l'on ne prend pas en compte dans les conclusions la part de surface vitrée bien différente en fonction des orientations comme on peut le voir sur la figure 4.5.

En ce qui concerne les pertes en figure 4.4a et 4.4b, on remarque qu'elles sont moindres au Nord. Cette observation est cohérente avec le triple vitrage installé sur cette orientation (les autres orientations sont équipées de double vitrage) et les valeurs de référence correspondent bien à la médiane de la dispersion et il n'y a logiquement pas de différence de pertes entre les deux zones.

Pour les gains en figure 4.4c et 4.4d, le constat est différent, car l'exposition au soleil des vitrages va dépendre de l'orientation. Au Nord et à l'Est les gains sont faibles, ce qui est cohérent avec le fait qu'au mois de janvier ces deux orientations sont majoritairement

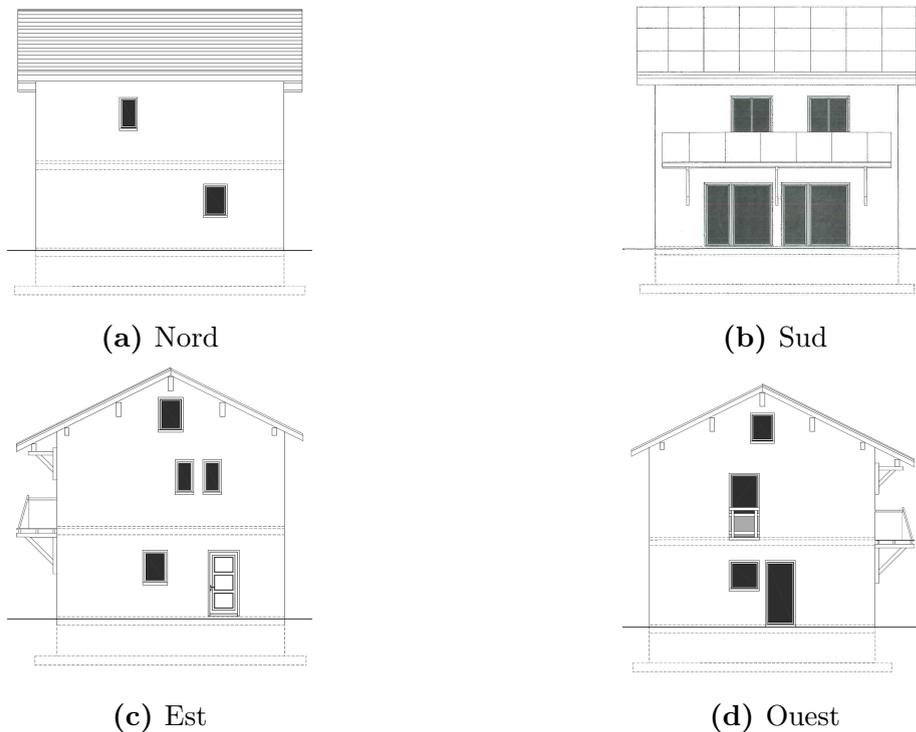


**Figure 4.4 – Lyon en janvier.** Dispersion sous forme de boxplot des pertes et gains par vitrage pour chaque zone et orientation.

- a) Pertes du rez-de-jardin
- b) Pertes de l'étage.
- c) Gains au rez-de-jardin
- d) Gains à l'étage.

à l'ombre (totalement pour le Nord et majoritairement pour l'Est, sachant que les fenêtres sont petites à l'Est et que les bords de fenêtre empêchent le soleil rasant). Ces conclusions ont été vérifiées à l'aide de Google Sketchup par la fonctionnalité "ombrage". Les gains au Sud sont importants pour les deux zones, mais supérieurs en zone rez-de-jardin ce qui peut s'expliquer par la présence à l'étage du balcon qui obstrue les surfaces vitrées du côté Sud. L'albédo au rez-de-jardin peut aussi jouer un rôle.

On observe, de manière globale que les pertes sont plus importantes que les gains, ce qui est justifié au mois de janvier pour le climat de Lyon. Les dispersions sont de l'ordre de 14% à 16% (toujours à 95% en niveau de confiance) quelque soit l'orientation et la zone, tandis que pour les orientations Sud et Ouest les variations sont de l'ordre de 32% à 36% sur les deux zones. L'étude de la dispersion montre une cohérence des échantillons



**Figure 4.5** – Distribution des surfaces vitrées de la maison INCAS-DM en fonction des orientations

avec le modèle physique en amont.

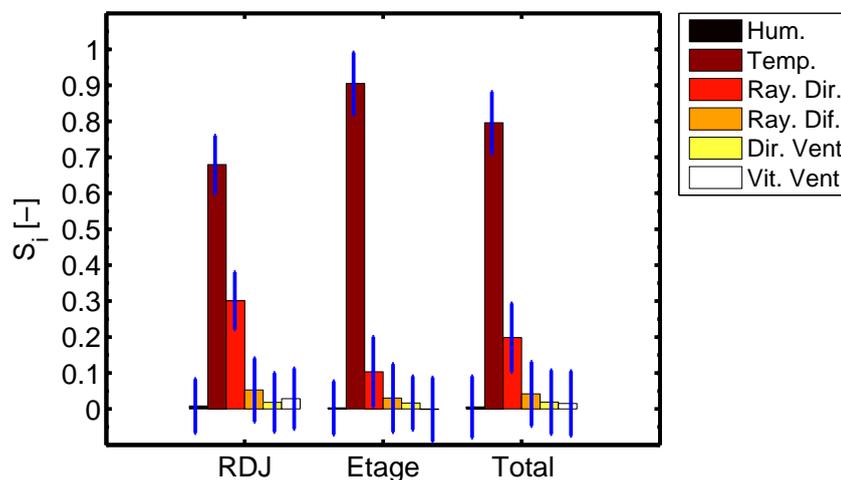
Afin de quantifier les effets de chaque variable météo sur les dispersions observées, on estime les indices de sensibilité.

### **Indice de sensibilité sur les besoins de chauffage : les variables responsables de la dispersion**

La figure 4.6 montre sous forme d'histogramme les valeurs de sensibilité des six variables étudiées sur les trois sorties de besoins de chauffage. Les bars d'erreurs sont obtenues par 100 répliques bootstrap, l'ordre de grandeur de l'écart type bootstrap est de 0,05 i.e. 10% à 95% de niveau de confiance. Nous voyons en figure 4.6 que 80% de l'incertitude de besoin de chauffage total est due à la variation de la température tandis que le rayonnement direct explique 20% de la variation. Nous remarquons que seul le rayonnement direct et la température sont influents et l'humidité relative, la direction et la vitesse de vent ne sont pas influents. Au fil des différents cas présentés dans ce chapitre d'application sur la maison INCAS nous verrons que ces variables météorologiques ne sont pas influentes sur les sorties que nous observons que ce soit les besoins ou bien les pertes et gains par vitrage, ce qui s'explique par les modèles simplifiés utilisés dans le cas d'étude. Ainsi le

Le système de chauffage ou de refroidissement est un système idéal et ne prend donc pas en compte l'hygrométrie et seules les infiltrations permettent un contact d'air avec l'extérieur puisqu'il n'y a pas de renouvellement d'air. Cette simplification de modèle est volontaire pour permettre la comparaison sachant que le manque de données météo observées pour ces grandeurs (humidité, vitesse et direction de vent, rayonnement diffus) ne permet pas de commenter la variabilité en amont de l'analyse. Le rayonnement diffus quant à lui peut avoir une influence dans certains cas et notamment sur les sorties liées au vitrage.

La décomposition des besoins totaux de chaque zone permet d'apprécier la différence d'influence du rayonnement direct sur les deux sorties. Près de 30% de la variation du besoin de chauffage du rez-de-jardin explique les incertitudes contre 10% à l'étage ce qui s'explique par le pourcentage de surface vitrée supérieur au RDJ par rapport à la zone Etage comme nous pouvons le voir en figure 4.5 et particulièrement pour les façades Sud et Ouest qui sont les plus ensoleillées en janvier. Cela peut expliquer la différence d'influence du rayonnement direct. Les gains au Sud sont plus importants au RDJ pour des questions d'obstruction du balcon et par un effet de l'albédo.



**Figure 4.6 – Lyon en janvier** Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance calculées par 100 répliques bootstrap. On observe une influence du rayonnement direct plus importante au rez-de-jardin qu'à l'étage. La température extérieure est essentiellement responsable des variations de besoins de chauffage.

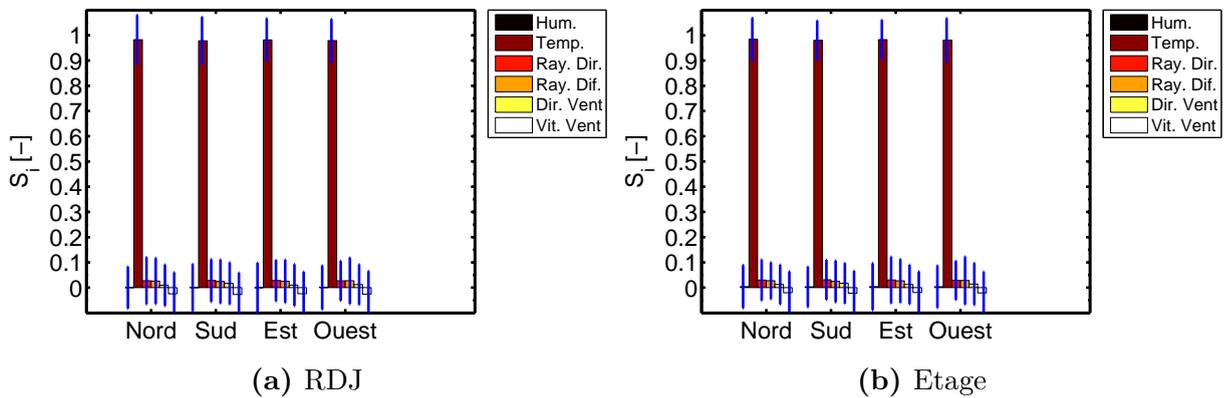
### Indices de sensibilité sur les gains et pertes par vitrage : vérification de la cohérence de la sensibilité

Nous étudions maintenant les indices de sensibilité associés à la variation des sorties pertes et gains par vitrage pour chaque orientation de chaque zone. L'étude des indices de sen-

sibilité permet de vérifier la cohérence de l'analyse sur des composants plus simples que le besoin de chauffage qui est une agrégation d'informations.

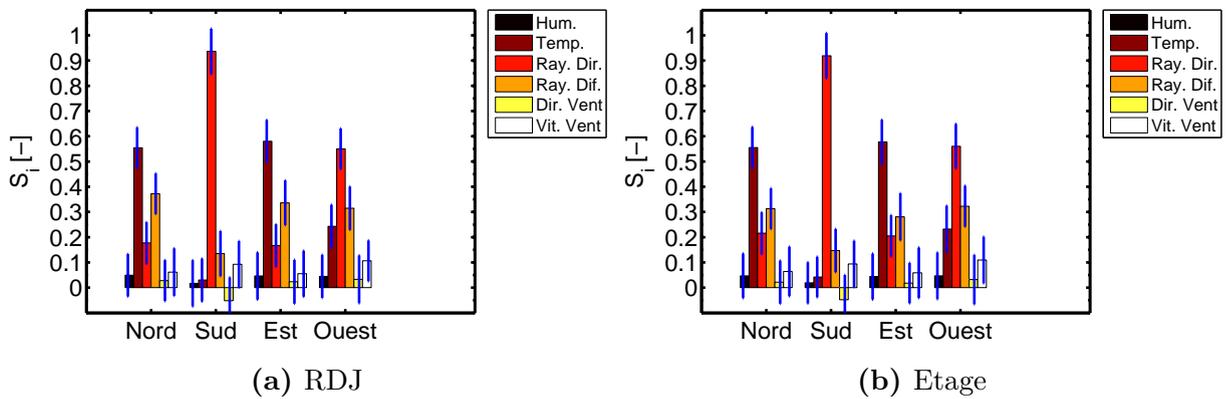
On observe en figure 4.7 que les variations sur les pertes ne sont expliquées que par la température ce qui est cohérent avec le climat du site et la période hivernale. Le bilan est plus mitigé pour les gains par vitrage représentés en figure 4.8 et les orientations peu ou pas exposées au soleil direct comme l'Est et le Nord ont une variabilité des pertes causée majoritairement par l'effet de la température qui représente environ 55% des variations des pertes au rez-de-jardin et à l'étage. Un peu plus de 30% des variations sont expliquées par le rayonnement diffus, le reste, beaucoup plus faible, est dû au rayonnement direct. Les indices de sensibilité sont cohérents avec le caractère ombragé de ces orientations et contraste avec l'orientation Sud où quasiment le seul paramètre responsable des variations sur les gains de vitrage est le rayonnement direct. La casquette du balcon est dimensionnée de manière à laisser pénétrer le soleil l'hiver car celui-ci est suffisamment bas et pour protéger du soleil l'été. L'orientation Ouest est ensoleillée en fin de journée, on observe donc une part des variations expliquées par le direct à plus de 50%, 30% par le diffus et les 20% restant par la température, ce qui est représentatif des différentes phases d'ensoleillement liées au cycle journalier du soleil.

La somme des indices de sensibilité est proche de 1, à la fois pour les gains et pertes mais aussi en ce qui concerne les besoins de chauffage. Nous pouvons en conclure que la totalité de l'information est présente et donc qu'il n'y a pas d'interactions entre variables dans le modèle. Ainsi, le calcul des effets principaux est suffisant et les effets totaux n'apporteraient pas d'informations supplémentaires.



**Figure 4.7 – Lyon en janvier.** Indice de sensibilité principal sur chaque orientations pour les pertes par les fenêtres de chaque zone. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance et obtenues par 100 répliques bootstrap. Les variations sur les pertes ne sont expliquées que par la température extérieure.

- a) Indice de sensibilité sur les pertes du rez-de-jardin
- b) Indice de sensibilité sur les pertes de l'étage



**Figure 4.8 – Lyon en janvier.** Indice de sensibilité principal sur chaque orientations pour les gains par les fenêtres de chaque zone. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance et obtenues par 100 répliques bootstrap. La sensibilité sur les gains dépend de l'exposition au soleil et donc de l'orientation.

- a) Indice de sensibilité sur les gains du rez-de-jardin
- b) Indice de sensibilité sur les gains de l'étage.

### Conclusion

Nous venons de quantifier l'incertitude associée à la variabilité naturelle de la météo pour le climat de Lyon au mois de janvier dans le cadre d'une maison basse consommation d'énergie. Nous avons montré que cette incertitude relative de plus de 20% sur le besoin

de chauffage total est due à la variation de la température. Dans le chapitre 3 nous avons caractérisé les variations des variables météorologiques qui sont représentatives de la variabilité des observations pour la température et l'ensoleillement dont la valeur moyenne est sous-évaluée. Il se peut qu'avec un fichier source IWEC aux valeurs d'ensoleillement plus proches de la réalité, que l'impact du rayonnement direct sur les incertitudes soit plus important, du moins pour le rez-de-jardin. L'étude de sortie plus simple et intuitive comme les gains et pertes par vitrage permet la vérification de la cohérence des indices de sensibilité obtenus ainsi que des dispersions. La somme des indices pour chaque sortie est proche de 1 car l'ensemble de l'information est présent dans les indices de sensibilité de premier ordre, il n'y a pas de perte d'information par interaction dans le modèle.

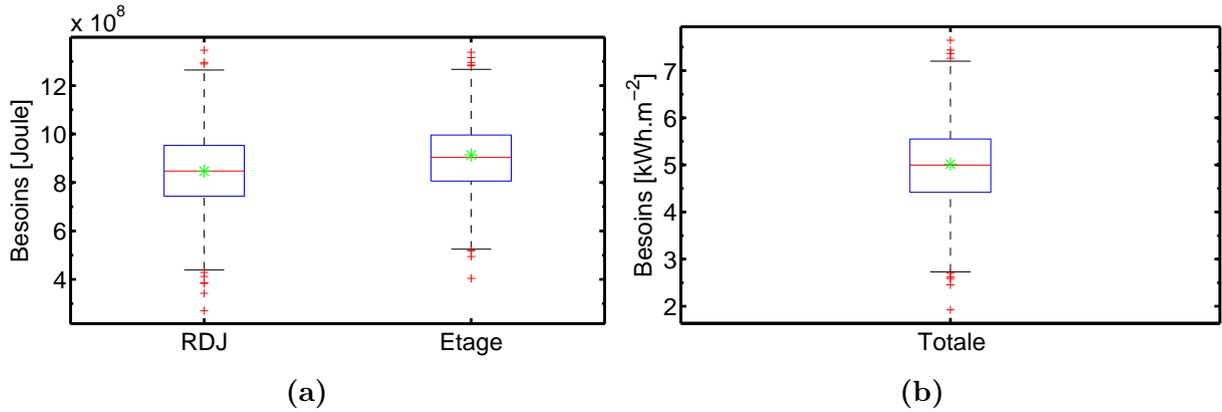
Nous avons vu aussi en chapitre 3 que l'échantillon généré pour le mois de janvier à Strasbourg présente des valeurs moyennes et de variabilité d'ensoleillement proches de l'échantillon lyonnais. Cependant les températures observées sont en moyenne plus froides de 36% par rapport à Lyon et la variabilité de l'échantillon est plus importante (elle englobait quasiment l'ensemble des observations à 95% de confiance). Quelles différences allons nous observer pour le climat de Strasbourg sur l'incertitude et concernant les indices de sensibilité ?

#### 4.2.1.2 Comparaison des résultats obtenus avec le climat de Strasbourg

##### Dispersion des résultats

Tout comme la description détaillée de Lyon au mois de janvier nous allons étudier la dispersion des pertes et gains par vitrage pour chaque orientation et chaque zone, ce qui nous permettra de vérifier le comportement du bâtiment et de le comparer à celui situé à Lyon. Afin d'alléger le propos, la dispersion sous forme de boxplot est présentée mais la comparaison à la normale n'est pas montrée. Les résultats sont résumés en un tableau récapitulatif avec les résultats de Lyon pour faciliter la comparaison avec le climat de Strasbourg.

Tout d'abord on observe que les valeurs de référence correspondent bien à la médiane de la dispersion empirique et que la dispersion est symétrique. A partir des valeurs des écarts types et des moyennes on obtient l'intervalle de confiance correspondant à un seuil de 95% soit  $1,96\sigma$ . Afin de visualiser les écarts, la variation en % par rapport à la valeur moyenne est aussi renseignée en considérant une probabilité de 95% d'obtenir cette valeur.



**Figure 4.9 – Strasbourg en janvier.** Dispersion des 2000 valeurs de besoins de chauffage. Représentation sous forme de boxplot avec l'étoile verte comme marque de la simulation de référence effectuée avec le IWECC source.

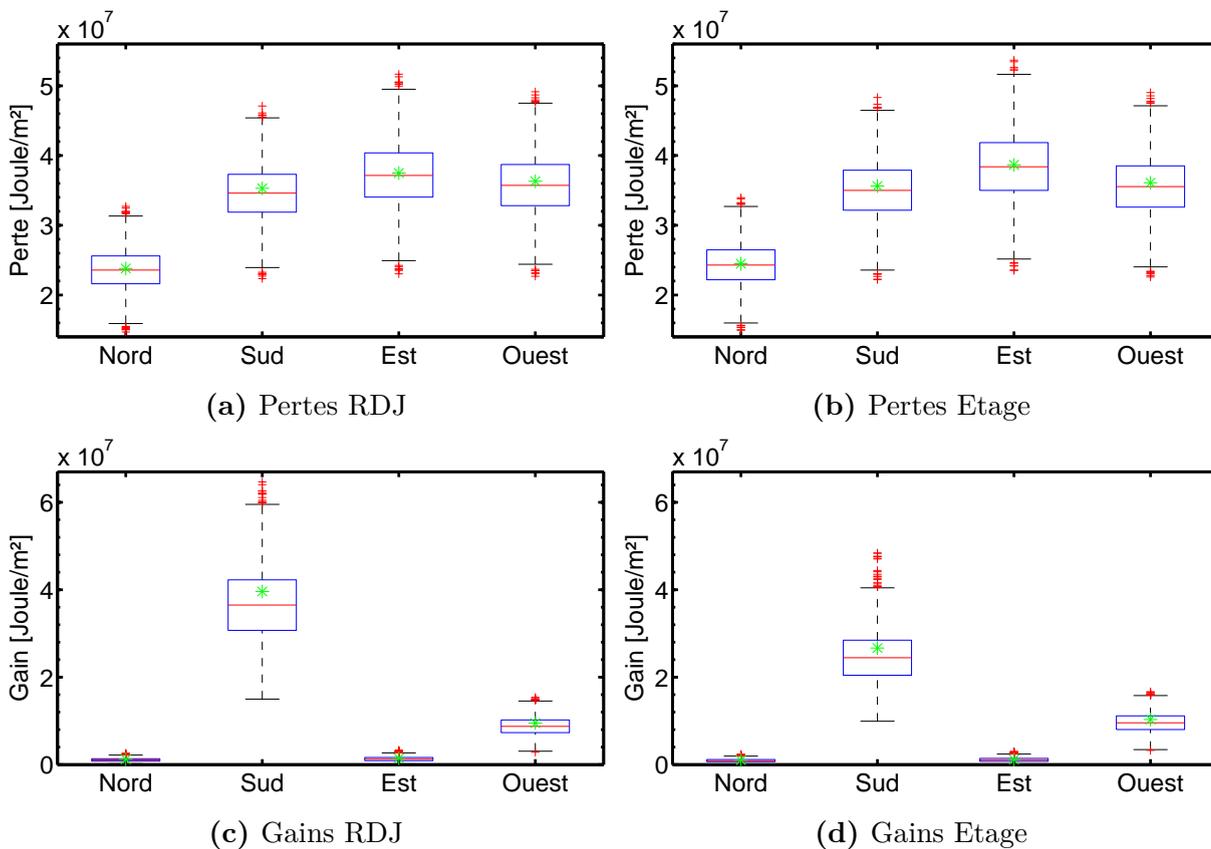
- a) Besoins sur les deux zones : le rez-de-jardin et l'étage
- b) Besoins de chauffage total

**Table 4.2 – Strasbourg en janvier** Valeurs de dispersion du besoin de chauffage total et par zone

		Résultats Strasbourg			
Chauffage		Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $10^8$ Joule]		8,46	1,54	$8,46 \pm 3,02$	35,6%
Etage [ $10^8$ Joule]		9,02	1,35	$9,02 \pm 2,65$	29,4%
Totale [ $kWh.m^{-2}$ ]		4,98	0,816	$4,98 \pm 1,60$	32,1%
		Récapitulatif Lyon			
Chauffage		Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $10^8$ Joule]		7,96	0,990	$7,96 \pm 1,94$	24,4%
Etage [ $10^8$ Joule]		8,36	0,841	$8,36 \pm 1,65$	19,7%
Totale [ $kWh.m^{-2}$ ]		4,65	0,516	$4,65 \pm 1,01$	21,7%

Le tableau 4.2 regroupe les résultats des dispersions des 2000 valeurs de besoins de chauffage pour Strasbourg. Le besoin de chauffage moyen est supérieur de 6% par rapport à Lyon, ce qui est cohérent avec des températures plus froides en hiver pour Strasbourg. On observe que la dispersion des résultats est plus importante à Strasbourg qu'à Lyon, ce qui est cohérent avec le fait que la dispersion des échantillons de température est elle aussi plus importante pour Strasbourg que pour Lyon.

De nouveau, afin de vérifier la dispersion sur une sortie simple, la dispersion des gains et pertes par vitrage pour chaque orientation et pour chaque zone est étudiée par boxplot.



**Figure 4.10 – Strasbourg en janvier.** Dispersion sous forme de boxplot des pertes et gains par vitrage pour chaque zone et orientation.

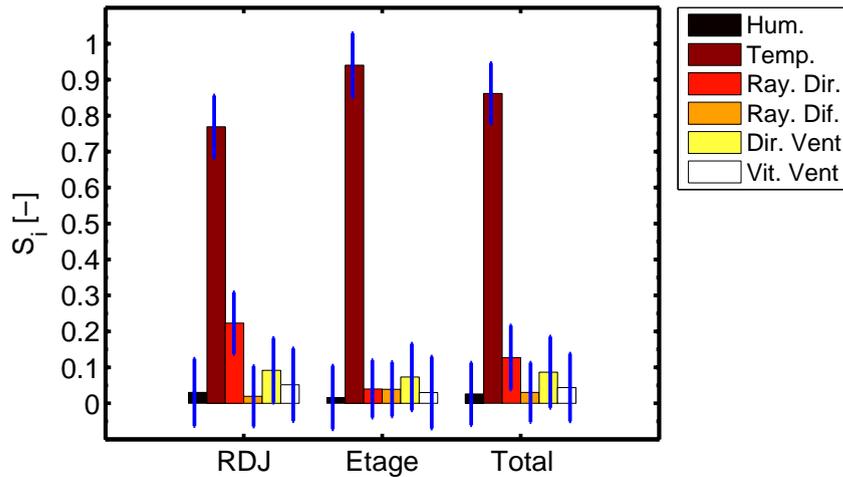
- a) Pertes du rez-de-jardin
- b) Pertes de l'étage
- c) Gains au rez-de-jardin
- d) Gains à l'étage

On observe sur la figure 4.10 le même ordre de grandeur pour les gains et pertes que le cas de Lyon en fonction des orientations. Les gains Nord et Est déjà peu élevés sont

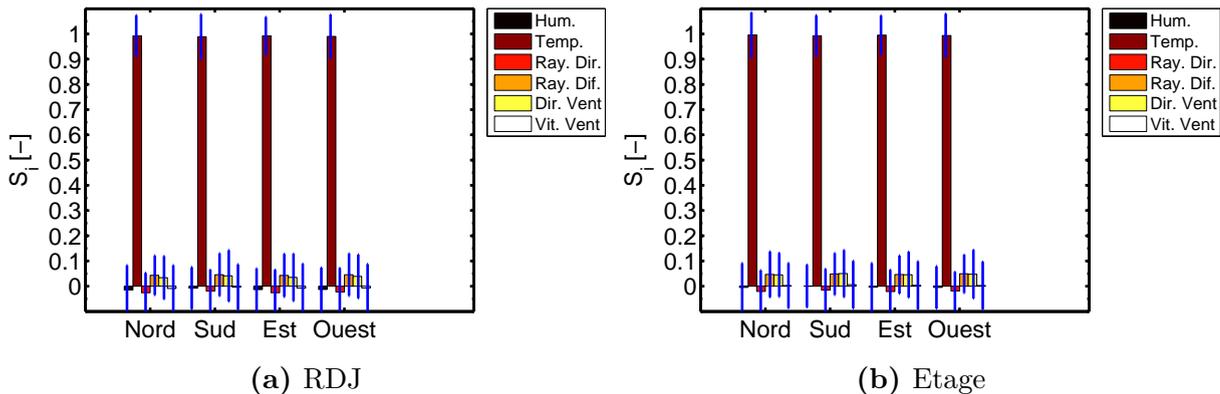
divisés par 2. La dispersion est plus importante : les pertes de 22% à 25% et les gains pour le Sud et l'Ouest où les gains sont les plus importants on observe une dispersion de 45% à 47%. Afin de quantifier les effets de chaque variable météo sur les dispersions observées, les indices de sensibilités sont estimés.

### **Indice de sensibilité**

La figure 4.11 représente l'influence des six variables météorologiques sur la dispersion des trois sorties de performance où 20 % de l'incertitude sur le RDJ est expliquée par le rayonnement direct, ce qui représente une réduction de 10% par rapport à Lyon. On constate la même baisse d'influence pour la zone Etage, ce qui est cohérent avec l'augmentation des besoins entre Strasbourg et Lyon. En hiver l'interprétation des effets de chaque variable influente est simple, le rayonnement direct diminue les besoins alors que la température les augmente. En été l'interprétation est plus complexe car la température a deux effets antagonistes en fonction de la journée, elle diminue les besoins de refroidissement la nuit et les augmente la journée.

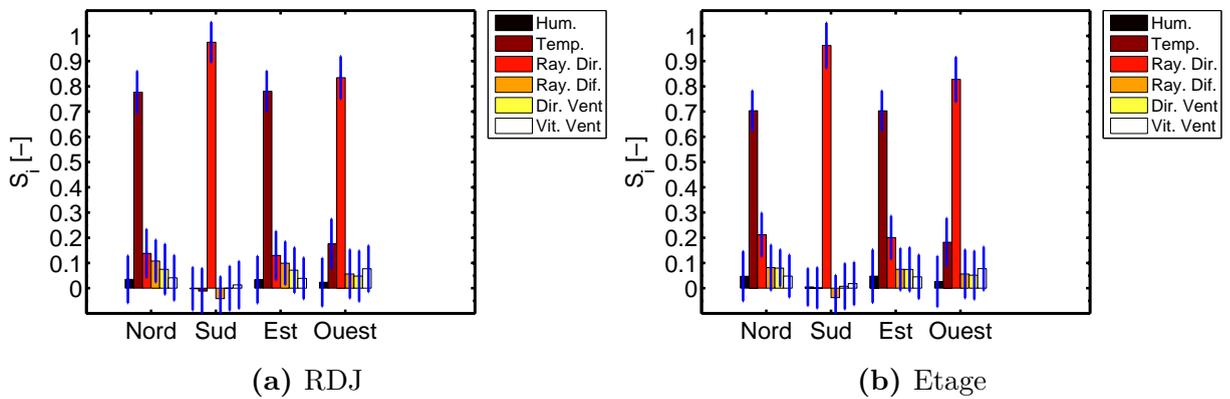


**Figure 4.11 – Strasbourg en janvier.** Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance et obtenues par 100 répliques bootstrap.



**Figure 4.12 – Strasbourg en janvier.** Indice de sensibilité principal sur chaque orientation pour les pertes par les fenêtres de chaque zone. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance et obtenues par 100 répliques bootstrap. Les variations sur les pertes ne sont expliquées que par la température extérieure.

- a) Indice de sensibilité sur les pertes du rez-de-jardin
- b) Indice de sensibilité sur les pertes de l'étage



**Figure 4.13 – Strasbourg en janvier.** Indice de sensibilité principal sur chaque orientation pour les gains par les fenêtres de chaque zone. Bar d’erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance et obtenues par 100 répliques bootstrap. Quelque soit l’orientation, le rayonnement diffus ne fait pas partie des paramètres influents contrairement au cas de Lyon.

- a) Indice de sensibilité sur les gains du rez-de-jardin
- b) indice de sensibilité sur les gains de l’étage

On observe en figure 4.12 et 4.13 la sensibilité des six variables sur la dispersion des seize sorties du composant fenêtre. La dispersion des pertes en figure 4.12 est expliquée exclusivement par la température. Pour les gains en figure 4.13, les températures de l'échantillon de Strasbourg sont plus basses et la variabilité plus importante, nous visualisons donc un indice de sensibilité associé plus élevé que pour Lyon dans le cas où les orientations n'entraînent pas d'exposition au rayonnement direct. On observe côté Ouest un effet plus important du rayonnement direct pour Strasbourg que pour Lyon car le rayonnement diffus n'a pas d'influence pour Strasbourg, et la variabilité et la valeur du rayonnement diffus journalier moyen sont moins élevées pour Strasbourg. Enfin, la somme des indices est proche de 1, il n'y a donc pas de perte d'informations.

#### 4.2.1.3 Conclusion

La réponse du bâtiment en période hivernale a été étudiée pour deux types d'échantillons qui ont des valeurs ainsi qu'une dispersion proches pour l'ensoleillement mais des températures plus variables pour Strasbourg et plus froides de 36% sur la valeur moyenne quotidienne. L'écart au besoin de chauffage total est de 6% mais la dispersion est plus importante et l'incertitude relative à 95% est de plus de 30% soit 10% de plus que pour le cas de Lyon. Les indices de sensibilité montrent que les besoins de chauffage sont plus sensibles à la température à Strasbourg qu'à Lyon.

### 4.2.2 Réponse du bâtiment en période estivale

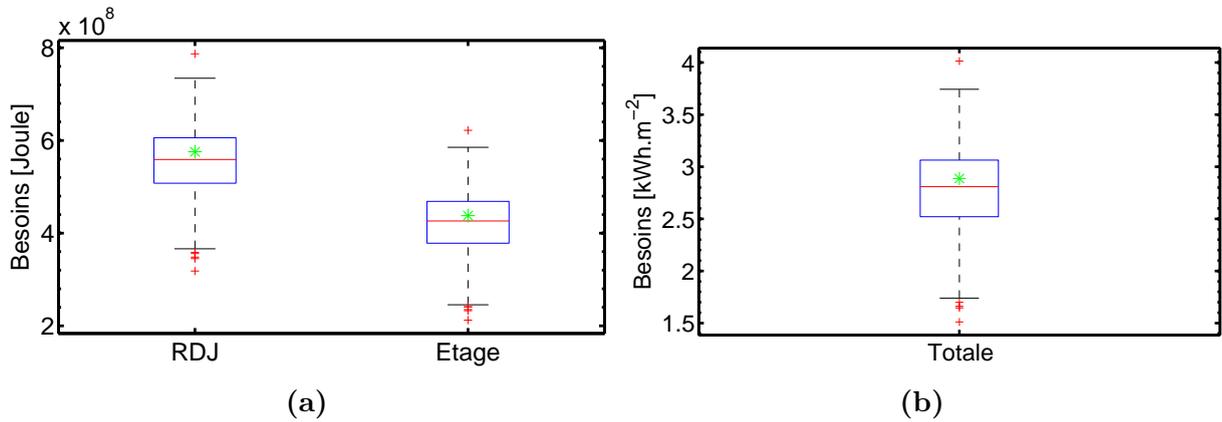
#### 4.2.2.1 Résultats obtenus pour le climat de Lyon

##### Dispersion des résultats

A partir des valeurs des écarts types et des moyennes on obtient l'intervalle de confiance correspondant à un seuil de 95% soit  $1,96\sigma$ . Pour mieux visualiser les écarts, la variation par rapport à la valeur moyenne est aussi renseignée en considérant une probabilité de 95% d'obtenir cette valeur.

**Table 4.3 – Lyon en juillet** Valeurs de dispersion du besoin de refroidissement total et par zone

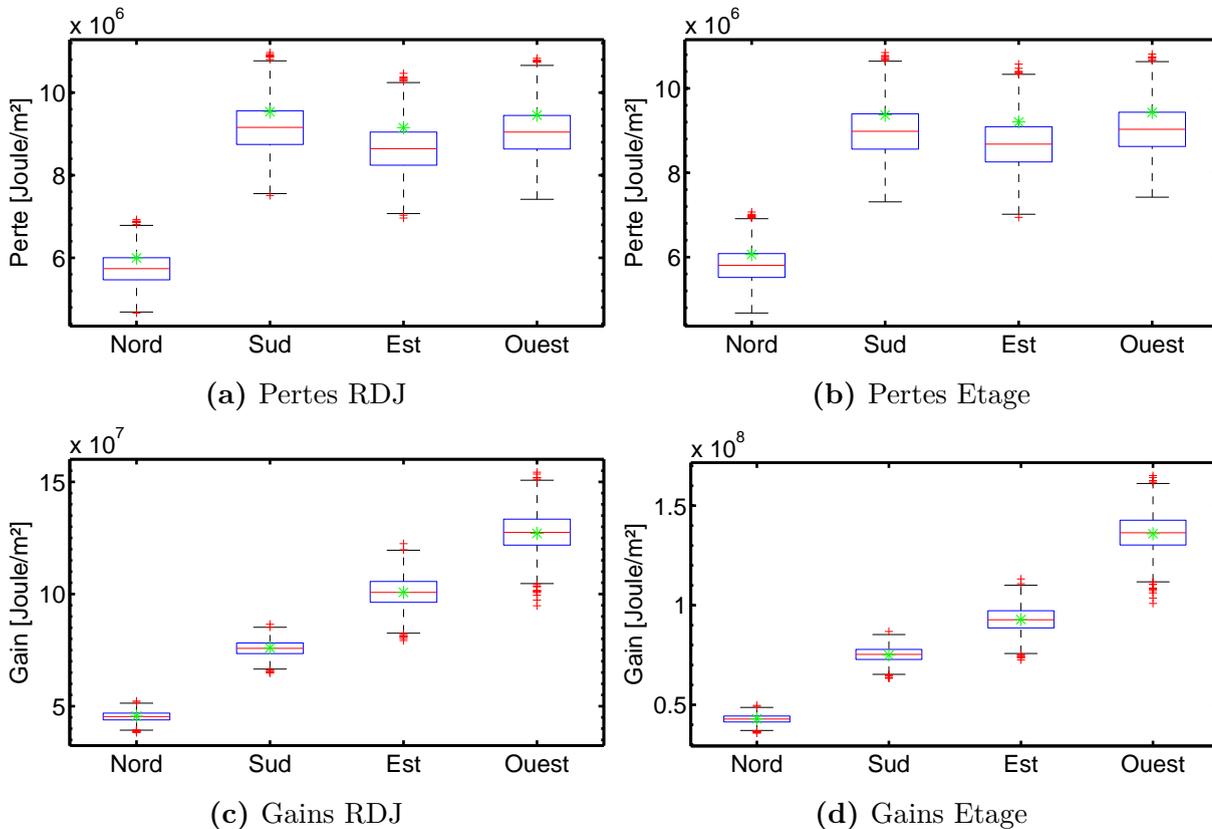
Refroidissement	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $.10^8$ Joule]	5,57	0,711	$5,57 \pm 1,39$	25,0%
Etage [ $.10^8$ Joule]	4,24	0,640	$4,24 \pm 1,25$	29,6%
Totale [ $kWh.m^{-2}$ ]	2,79	0,384	$2,79 \pm 0,75$	27,0%



**Figure 4.14 – Lyon en juillet.** Dispersion des 2000 valeurs de besoins de refroidissement. Représentation sous forme de boxplot avec l'étoile verte comme marque de la simulation de référence effectuée avec le IWEC source.

- a) Besoins pour les deux zones : le rez-de-jardin et l'étage
- b) Besoin de chauffage total

Le tableau 4.3 regroupe les caractéristiques des dispersions pour les besoins de refroidissement totaux et de chaque zone et l'écart à la moyenne peut atteindre 27% à 95% de seuil de confiance. Nous avons observé un besoin de chauffage plus important en hiver (cf. tableau 4.1), à l'étage, montrant que les températures de l'étage sont plus froides. Dans le cas de besoin de refroidissement, on observe également un écart, cette fois plus important, entre rez-de-jardin et étage, avec un besoin de refroidissement supérieur au rez-de-jardin.



**Figure 4.15 – Lyon en juillet.** Dispersion sous forme de boxplot des pertes et gains par vitrage pour chaque zone et orientation.

- a) Pertes du rez-de-jardin
- b) Pertes de l'étage
- c) Gains au rez-de-jardin
- d) Gains à l'étage.

La figure 4.15 regroupe les dispersions pour chaque orientation des pertes et gains par fenêtre sur chaque zone et de manière générale les pertes sont faibles et les gains plus importants que pour le cas hivernal. Les pertes au Nord sont plus faibles que pour les autres orientations ce qui est dû au triple vitrage. On a la même hiérarchisation des pertes mais pour les gains l'ensoleillement en fonction des orientations n'est pas le même et les casquettes au Sud sont efficaces. La surface Nord n'est au soleil qu'en fin de journée et la taille des fenêtres et les importants débords de fenêtre (35 cm) augmentent l'ombrage, ce qui explique le peu de gain. La casquette du sud protège des apports directs et limite les gains. L'Est est à l'ombre l'après-midi alors que l'Ouest n'est exposé au soleil que dans l'après-midi au moment où le rayonnement est le plus élevé. Ces différences d'exposition au cours de la journée expliquent les différents gains observés en figure 4.15. On note la bonne gestion de la casquette solaire pour le Sud, mais un éventuel manque de casquette

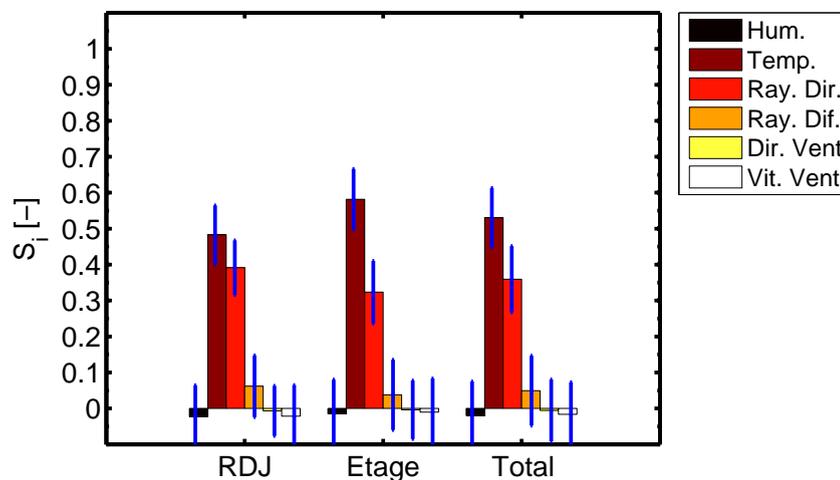
pour l'Ouest.

Les dispersions sont de l'ordre de 13% à 14% (toujours à 95% en niveau de confiance) pour les pertes et de 9% à 14% pour les gains.

Afin de quantifier les effets de chaque variable météo sur les dispersions observées et de vérifier les interprétations énoncées ci-dessus, on estime les indices de sensibilité.

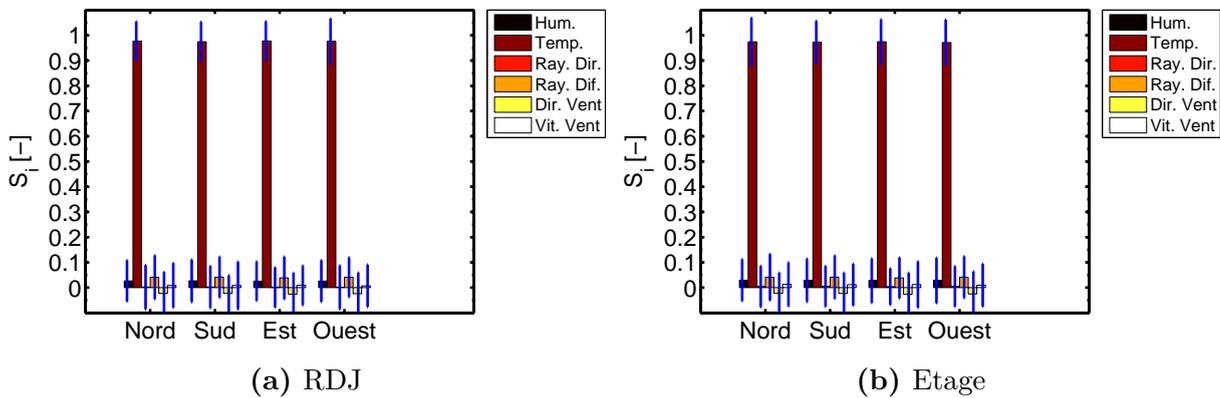
### Indices de sensibilité sur les dispersions

Selon les indices de sensibilité pour le cas de Lyon au mois de juillet en figure 4.16, 30% à 40% de la dispersion sur les besoins de refroidissement sont dus au rayonnement solaire direct. Encore une fois nous observons que le rayonnement solaire direct est plus influent au RDJ qu'à l'Etage, cela explique en partie la raison pour laquelle les besoins de refroidissement sont plus importants au RDJ. Le rayonnement solaire va contribuer à l'augmentation du besoin de refroidissement, la température par contre va contribuer la journée à une augmentation des besoins mais la nuit elle peut contribuer à la diminution de ces besoins en refroidissement.



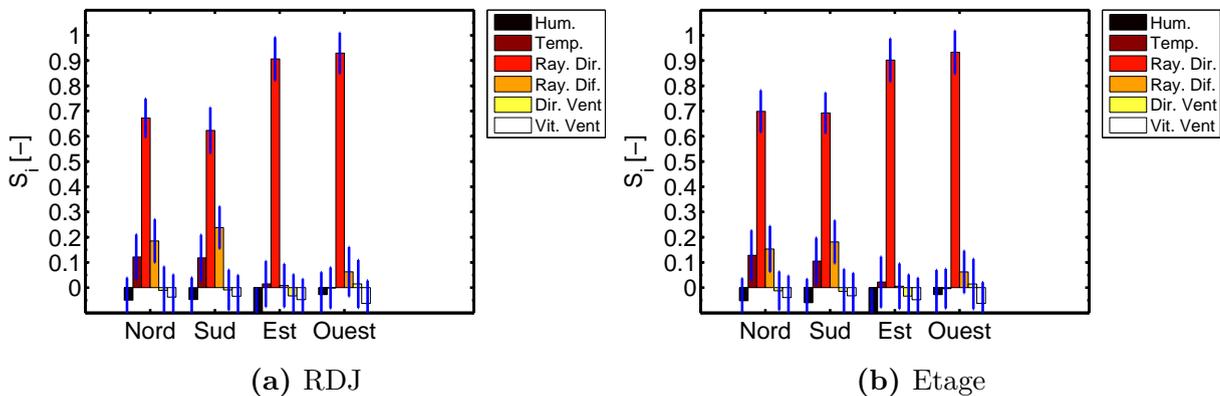
**Figure 4.16 – Lyon en juillet** Indice de sensibilité principal sur le besoin de refroidissement des deux zones et le besoin total. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance calculé par 100 répliques bootstrap. On observe une influence du rayonnement direct plus importante au rez-de-jardin qu'à l'étage.

Les pertes ne sont pas influencées par le rayonnement car elles se produisent la nuit, le rayonnement solaire ne peut donc pas venir limiter les pertes quand celles-ci se produisent.



**Figure 4.17 – Lyon en juillet.** Indice de sensibilité principal sur chaque orientation pour les pertes par les fenêtres de chaque zone. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance et obtenues par 100 répliques bootstrap. Les variations des pertes dépendent exclusivement de la variation de la température.

- a) Indice de sensibilité sur les pertes du rez-de-jardin
- b) Indice de sensibilité sur les pertes de l'étage



**Figure 4.18 – Lyon en juillet.** Indice de sensibilité principal sur chaque orientations pour les gains par les fenêtres de chaque zone. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance et obtenues par 100 répliques bootstrap.

- a) Indice de sensibilité sur les gains du rez-de-jardin
- b) Indice de sensibilité sur les gains de l'étage

Pour les gains on retrouve deux types de comportements, soit le vitrage est ombragé et la variation est due à une part de diffus comme au Sud et au Nord, soit le soleil pénètre dans la zone par le vitrage comme pour les orientations Ouest l'après midi et Est le matin où dans ce cas le rayonnement direct domine les variations de sortie.

4.2.2.2 Résultats obtenus pour le climat de Strasbourg

Les échantillons de Strasbourg et de Lyon disposent de la même variabilité au mois de juillet mais la valeur moyenne d'ensoleillement pour Strasbourg est de 23% plus faible. De même les températures sont moins élevés de 15% pour Strasbourg par rapport à Lyon ce qui entraîne des besoins de refroidissement plus faibles. Les deux échantillons ayant les mêmes variabilités, les conclusions seront plus aisées et seulement basées sur le décalage de valeur moyenne.

Dispersion des résultats

A partir des valeurs des écarts types et des moyennes, on obtient l'intervalle de confiance correspondant à un seuil de 95% soit  $1,96\sigma$ . Pour mieux visualiser les écarts, la variation par rapport à la valeur moyenne est aussi renseignée, en considérant une probabilité de 95% d'obtenir cette valeur.

Le tableau 4.4 regroupe les résultats des dispersions des 2000 valeurs de besoins de chauffage pour Strasbourg. Le besoin de refroidissement moyen est inférieur de 44% par rapport à Lyon, ce qui est cohérent avec des températures plus fraîches en été à Strasbourg. On observe le même écart type pour Lyon et Strasbourg mais comme les besoins sont moins importants la dispersion relative des résultats est plus importante à Strasbourg qu'à Lyon.

Table 4.4 – Strasbourg en juillet Valeurs de dispersion du besoin de refroidissement total et par zone

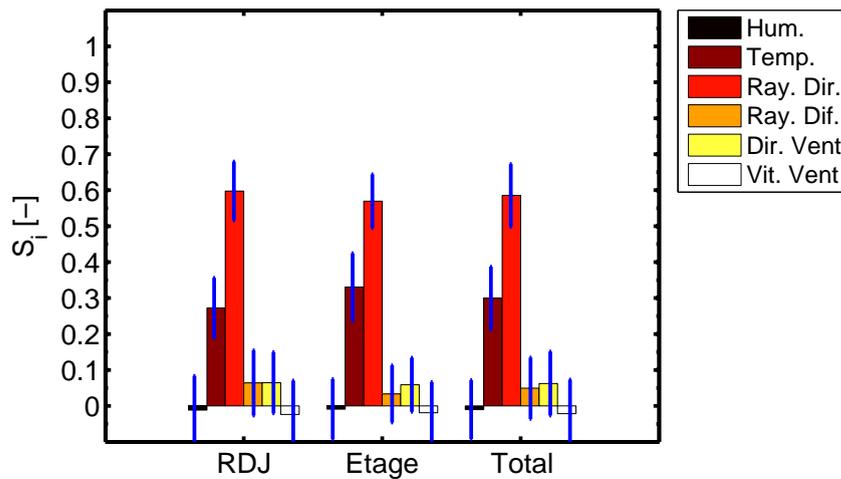
	Résultats Strasbourg			
Refroidissement	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $\cdot 10^8$ Joule]	3,24	0,741	$3,24 \pm 1,45$	44,8%
Etage [ $\cdot 10^8$ Joule]	2,29	0,648	$2,29 \pm 1,27$	55,5%
Totale [ $kWh.m^{-2}$ ]	1,57	0,395	$1,57 \pm 0,78$	49,2%
	Récapitulatif Lyon			
Refroidissement	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $\cdot 10^8$ Joule]	5,57	0,711	$5,57 \pm 1,39$	25,0%
Etage [ $\cdot 10^8$ Joule]	4,24	0,640	$4,24 \pm 1,25$	29,6%
Totale [ $kWh.m^{-2}$ ]	2,79	0,384	$2,79 \pm 0,75$	27,0%

En ce qui concerne les fenêtres, les valeurs sont plus faibles pour les gains et plus élevés pour les pertes, ce qui est cohérent avec les échantillons météo de Strasbourg. La dispersion est de l'ordre de 10% pour les pertes et de 15% pour les gains ce qui est le

même ordre de grandeur de dispersion que pour Lyon.

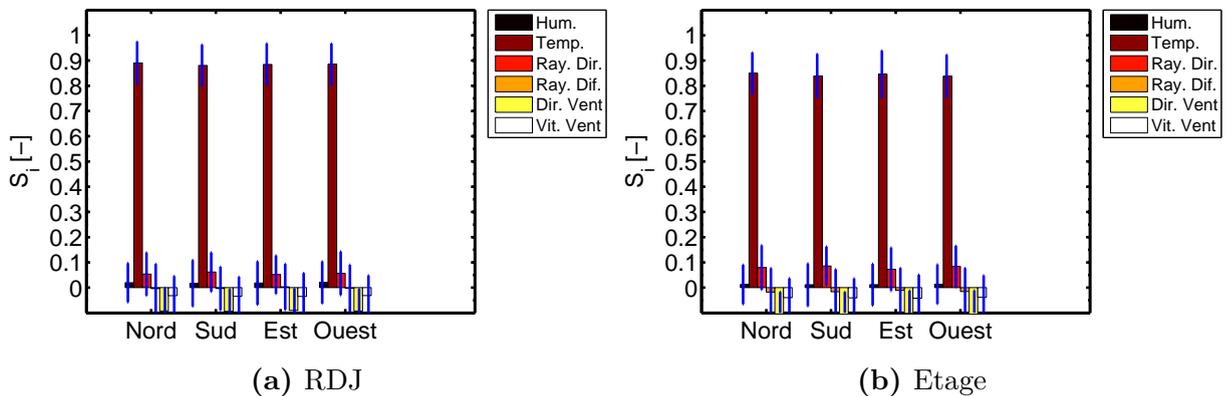
### Indice de sensibilité

Nous avons la même variabilité donc si on observe des différences de sensibilité, cela est provoqué par le décalage de la valeur nominale et en d'autres termes de l'espace que nous explorons.



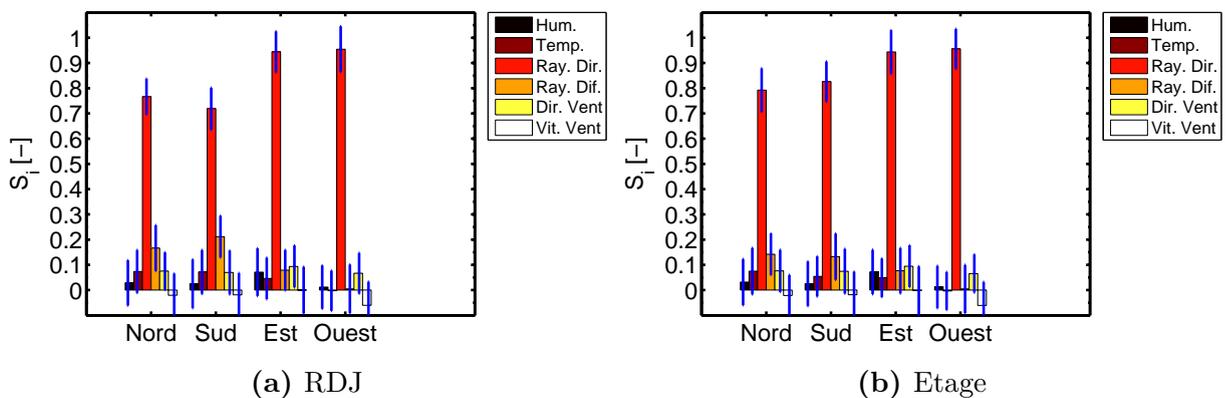
**Figure 4.19 – Strasbourg en juillet** Indice de sensibilité principal sur le besoin de refroidissement des deux zones et le besoin total. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance calculé par 100 répliques bootstrap.

L'influence du rayonnement direct est plus importante à Strasbourg qu'à Lyon bien que les valeurs d'ensoleillement soient moins importantes à Strasbourg. Cela s'explique par le fait que les valeurs de températures étant plus faibles à Strasbourg, le besoin de refroidissement va davantage dépendre du rayonnement solaire qu'à Lyon où les températures contribuent à l'augmentation des besoins de refroidissement.



**Figure 4.20 – Strasbourg en juillet.** Indice de sensibilité principal sur chaque orientations pour les pertes par les fenêtres de chaque zone. Bar d’erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance et obtenues par 100 répliques bootstrap. La variation de sortie est expliquée exclusivement par la température extérieure

- a) Indice de sensibilité sur les gains du rez-de-jardin
- b) Indice de sensibilité sur les gains de l’étage.



**Figure 4.21 – Strasbourg en juillet.** Indice de sensibilité principal sur chaque orientations pour les gains par les fenêtres de chaque zone. Bar d’erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance et obtenues par 100 répliques bootstrap. Nous obtenons les mêmes résultats de sensibilité que pour Lyon

- a) Indice de sensibilité sur les gains du rez-de-jardin
- b) Indice de sensibilité sur les gains de l’étage

### 4.2.3 Conclusion

Nous avons montré que la dispersion, ainsi que la sensibilité obtenues avec deux types d’échantillons était cohérente avec la physique du bâtiment. En fonction de la variabilité et de la valeur de chaque variable on obtient des indices de sensibilité en accord. L’interprétation est plus complexe en été car la variable météo joue un double rôle : de refroidissement

de zone la nuit et de réchauffement la journée, ce qui rend les interprétations plus délicates. Nous avons donc réussi à évaluer une incertitude associée à la variabilité météo pour deux climats et à déterminer quelle variation d'entrée et donc quelle(s) variable(s) est (sont) responsable(s) de cette incertitude. Les résultats dépendent évidemment de la variabilité induite et du domaine d'exploration : climat chaud, climat froid. La validation indirecte de la méthodologie étant effectuée, nous proposons maintenant de comparer sur un même climat (Lyon) des cas technologiques en modélisation différents.

## 4.3 Comparaison de modèles pour le même climat

### 4.3.1 Résultats pour le cas "Triple vitrage"

#### 4.3.1.1 Protocole

La modèle thermique test de la maison INCAS-DM que nous utilisons est constitué de fenêtres triple vitrage côté Nord et de fenêtres double vitrage sur le reste des orientations. Le triple vitrage permet une isolation plus importante du bâtiment, sa conductance est plus faible que le double vitrage. La façade Nord est la moins exposée au soleil en hiver, ce qui explique cette différence d'équipement. Cependant un triple vitrage est également moins transparent au rayonnement solaire, le facteur solaire du triple vitrage est inférieur à celui du double. Les valeurs de facteur solaire et de conductance pour le triple vitrage sont respectivement de 0,45 et 1,1  $W.m^{-2}.K^{-1}$  contre 0,6 et 1,4  $W.m^{-2}.K^{-1}$  pour le double vitrage.

Dans la partie suivante, nous souhaitons visualiser l'impact de la modification du double vitrage en triple pour les autres fenêtres sur la gestion des variables météorologiques par la maison INCAS-DM. Allons-nous observer une dispersion sur les besoins plus importante ? Une modification des variables qui en sont responsables ? Pour répondre à ces questions, nous estimons la dispersion sur les besoins en été et hiver de la maison INCAS-DM équipée de triple vitrage.

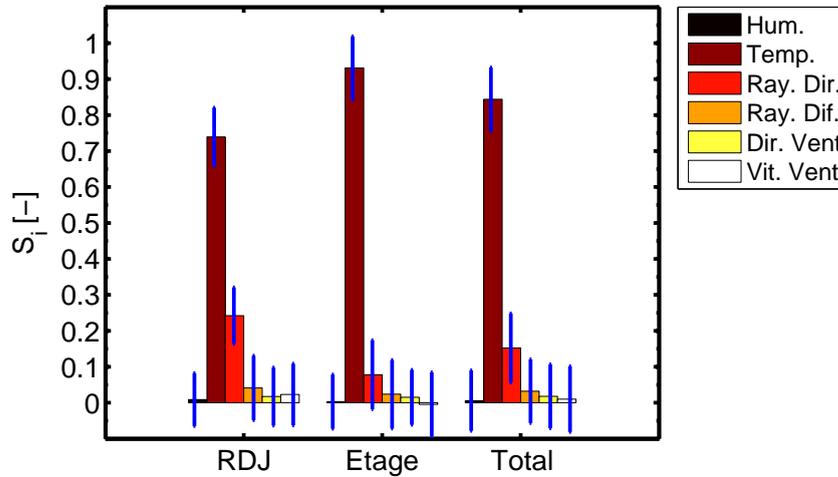
#### 4.3.1.2 Estimation en période hivernale

A partir des valeurs des écarts types et des moyennes on obtient l'intervalle de confiance correspondant à un seuil de 95% soit  $1,96\sigma$ . Pour mieux visualiser les écarts, la variation par rapport à la valeur moyenne est aussi renseignée en considérant une probabilité de 95% d'obtenir cette valeur.

**Table 4.5 – Triple vitrage en janvier** Valeurs de dispersion du besoin de chauffage total et par zone

	<b>Résultats Triple Vitrage</b>			
<b>Chauffage</b>	<b>Moyenne <math>\mu</math></b>	<b>Ecart type <math>\sigma</math></b>	<b>Intervalle de confiance</b>	<b>Variation</b>
RDJ [ $.10^8$ Joule]	7,53	0,839	$7,53 \pm 1,64$	21,8%
Etage [ $.10^8$ Joule]	7,98	0,766	$7,98 \pm 1,50$	18,8%
Totale [ $kWh.m^{-2}$ ]	4,42	0,453	$4,42 \pm 0,89$	20,1%
	<b>Récapitulatif cas test original</b>			
<b>Chauffage</b>	<b>Moyenne <math>\mu</math></b>	<b>Ecart type <math>\sigma</math></b>	<b>Intervalle de confiance</b>	<b>Variation</b>
RDJ [ $.10^8$ Joule]	7,96	0,990	$7,96 \pm 1,94$	24,4%
Etage [ $.10^8$ Joule]	8,36	0,841	$8,36 \pm 1,65$	19,7%
Totale [ $kWh.m^{-2}$ ]	4,65	0,516	$4,65 \pm 1,01$	21,7%

Selon les résultats du tableau 4.5, le triple vitrage apporte une diminution de 5% des besoins de chauffage par rapport au cas test original. La dispersion de sortie est du même ordre de grandeur dans le cas du triple vitrage et du double. Si on étudie les pertes et gains par vitrage, on détermine une diminution des pertes d'environ 35% et des gains allant jusqu'à 25% pour les orientations les plus exposées comme le Sud et l'Ouest. Est ce que ces modifications de pertes et de gains vont impacter les variables responsables de la dispersion sur les besoins? Pour y répondre, les indices de sensibilité sont estimés et présentés sous forme d'histogramme en figure 4.22.



**Figure 4.22 – Triple Vitrage en janvier** Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance calculé par 100 répliques bootstrap. On observe une influence du rayonnement direct plus importante au rez-de-jardin qu'à l'étage. La température extérieure est essentiellement responsable des variations de besoin de chauffage.

Les valeurs des effets principaux  $S_i$  ainsi que la moyenne du bootstrap  $\hat{S}_i$  sont regroupées dans le tableau 4.6. Les valeurs de sensibilité pour le cas "Triple Vitrage" sont du même ordre de grandeur que celles observées pour le cas test original. La température explique majoritairement la dispersion de sortie. On observe une légère baisse de l'effet du rayonnement direct dans le cas du "Triple Vitrage" notamment au rez-de-jardin. Le rez-de-jardin est une zone au pourcentage de surface vitrée plus élevé que la zone Etage et notamment du côté Sud. Cette configuration lui confère une sensibilité au solaire plus importante qu'à l'Etage comme l'atteste la sensibilité du rayonnement direct pour le cas test original. L'écart de sensibilité entre les deux cas pourrait s'expliquer par la diminution du facteur solaire qui rend le bâtiment moins sensible au rayonnement et comme le rayonnement direct et la température extérieure sont les seuls paramètres influents, la sensibilité de la température augmente lorsque celle du rayonnement diminue. On garde cependant la même configuration de gestion de l'environnement par le bâtiment quelque soit le type de vitrage.

**Table 4.6 – Triple Vitrage en janvier** Valeurs des indices de sensibilité  $S_i$  et de son estimé bootstrap  $\hat{S}_i$  pour le besoin de chauffage total et de chaque zone

	Résultats Triple Vitrage					
	RDJ		Etage		Total	
	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$
$H_R$	0,009	-0,027	0,003	-0,027	0,006	-0,025
$T$	<b>0,739</b>	<b>0,731</b>	<b>0,931</b>	<b>0,919</b>	<b>0,844</b>	<b>0,840</b>
$R_{Dir}$	<b>0,242</b>	<b>0,214</b>	<b>0,078</b>	<b>0,057</b>	<b>0,152</b>	<b>0,121</b>
$R_{Diff}$	0,041	0,009	0,024	0,013	0,032	0,004
$V_{Dir}$	0,017	-0,013	0,015	-0,006	0,018	-0,005
$V_{Vit}$	0,023	-0,003	-0,005	-0,018	0,010	-0,007
	Récapitulatif cas test original					
	RDJ		Etage		Total	
	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$
$H_R$	0,008	-0,028	0,003	-0,029	0,006	-0,028
$T$	<b>0,680</b>	<b>0,667</b>	<b>0,905</b>	<b>0,892</b>	<b>0,796</b>	<b>0,790</b>
$R_{Dir}$	<b>0,301</b>	<b>0,274</b>	<b>0,103</b>	<b>0,081</b>	<b>0,199</b>	<b>0,168</b>
$R_{Diff}$	0,053	0,019	0,030	0,017	0,042	0,011
$V_{Dir}$	0,018	-0,013	0,0164	-0,006	0,019	-0,006
$V_{Vit}$	0,029	0,000	-0,001	-0,017	0,015	-0,003

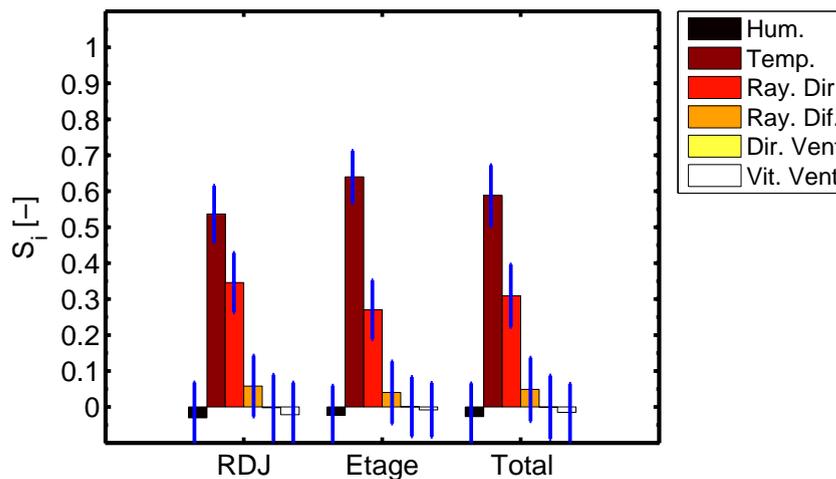
#### 4.3.1.3 Estimation en période estivale

A partir des valeurs des écarts types et des moyennes, l'intervalle de confiance correspondant à un seuil de 95% soit  $1,96\sigma$  est obtenu. Pour mieux visualiser les écarts, la variation en % par rapport à la valeur moyenne est aussi renseignée, toujours en considérant une probabilité de 95% d'obtenir cette valeur.

**Table 4.7 – Triple vitrage en juillet** Valeurs de dispersion du besoin de chauffage total et par zone

Résultats Triple Vitrage				
Refroidissement	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $.10^8$ Joule]	3,56	0,614	$3,56 \pm 1,20$	33,8%
Etage [ $.10^8$ Joule]	2,90	0,579	$2,90 \pm 1,13$	39,1%
Totale [ $kWh.m^{-2}$ ]	1,84,	0,339	$1,84 \pm 0,66$	36,1%
Récapitulatif cas test original				
Refroidissement	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $.10^8$ Joule]	5,57	0,711	$5,57 \pm 1,39$	25%
Etage [ $.10^8$ Joule]	4,24	0,640	$4,24 \pm 1,25$	29,6%
Totale [ $kWh.m^{-2}$ ]	2,79	0,384	$2,79 \pm 0,75$	27,0%

Selon les résultats du tableau 4.7, le triple vitrage apporte une diminution de 34% des besoins de refroidissement totaux et jusqu'à 36% pour la zone RDJ contre 32% pour l'Etage par rapport au cas test original. La dispersion de sortie est du même ordre de grandeur mais puisque la valeur moyenne diminue, l'incertitude relative est plus élevée de 10% dans le cas du triple vitrage. Si on étudie les pertes et gains par vitrage, on détermine une diminution des pertes d'environ 37% et des gains de 30%. Les indices de sensibilité sont estimés et présentés sous forme d'histogramme en figure 4.23.



**Figure 4.23 – Triple Vitrage en juillet** Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance calculé par 100 répliques bootstrap.

Les valeurs des effets principaux  $S_i$  ainsi que l'estimé bootstrap  $\hat{S}_i$  sont regroupées dans le tableau 4.8. Les valeurs de sensibilité pour le cas "Triple Vitrage" sont du même ordre

de grandeur que celles observées pour le cas test original. Comme dans le cas de la période hivernale, on observe une légère baisse de l'effet du rayonnement direct dans le cas "Triple Vitrage" mais cette diminution est observée sur les deux zones.

**Table 4.8 – Triple Vitrage en juillet** Valeurs des indices de sensibilité  $S_i$  et de son estimé bootstrap  $\hat{S}_i$  pour le besoin de chauffage total et de chaque zone

Résultats Triple Vitrage						
	RDJ		Etage		Total	
	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$
$H_R$	-0,030	-0,014	-0,023	-0,010	-0,027	-0,010
$T$	<b>0,537</b>	<b>0,557</b>	<b>0,640</b>	<b>0,655</b>	<b>0,588</b>	<b>0,599</b>
$R_{Dir}$	<b>0,346</b>	<b>0,360</b>	<b>0,270</b>	<b>0,291</b>	<b>0,309</b>	<b>0,322</b>
$R_{Diff}$	0,058	0,078	0,040	0,059	0,049	0,064
$V_{Dir}$	-0,002	0,013	0,001	0,015	-0,001	0,017
$V_{Vit}$	-0,022	-0,003	-0,008	0,008	-0,015	-0,002
Récapitulatif cas test original						
	RDJ		Etage		Total	
	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$
$H_R$	-0,023	-0,009	-0,015	-0,006	-0,019	-0,007
$T$	<b>0,483</b>	<b>0,490</b>	<b>0,582</b>	<b>0,591</b>	<b>0,530</b>	<b>0,530</b>
$R_{Dir}$	<b>0,391</b>	<b>0,404</b>	<b>0,323</b>	<b>0,334</b>	<b>0,359</b>	<b>0,377</b>
$R_{Diff}$	0,062	0,070	0,038	0,045	0,049	0,054
$V_{Dir}$	-0,007	0,018	-0,004	0,022	-0,006	0,020
$V_{Vit}$	-0,021	-0,002	-0,010	0,003	-0,016	0,002

En comparant le modèle original de la maison DM avec un modèle où les vitrages ont été remplacés par du triple, on observe sur les deux périodes une légère diminution de l'effet du rayonnement solaire sur le bâtiment. Dans le cadre du modèle avec triple vitrage la dispersion de sortie est davantage liée à la variabilité de la température.

### 4.3.2 Résultats pour un bâtiment moins performant : cas "Maison Dégradée"

#### 4.3.2.1 Protocole

On souhaite évaluer la dispersion due à la variabilité météo sur un bâtiment moins performant que le cas test original qu'est la maison DM en dégradant virtuellement les performances du modèle thermique. Les infiltrations des zones RDJ et Etage sont multipliés

par 2, les épaisseurs d'isolant sont divisées par 2 et une épaisseur de mur est supprimée pour les mur extérieurs ainsi que pour le vide sanitaire (voir annexes pour les coupes de principe et le schéma du double-mur). Les modifications au niveau des murs extérieurs engendrent la suppression des casquettes de bords de fenêtre ce qui limite les ombrages. Le modèle thermique obtenu reste optimisé et loin d'un bâtiment énergivore et nous allons évaluer la dispersion de la variabilité météo sur ce modèle dégradé.

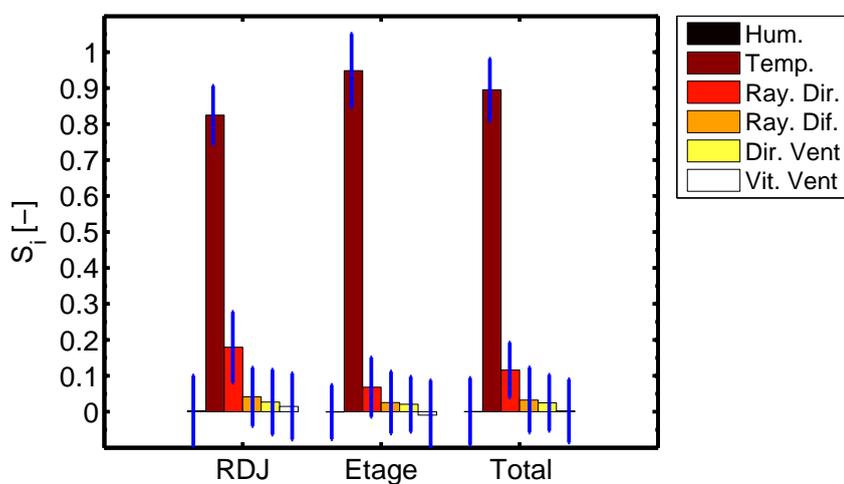
#### 4.3.2.2 Résultats de l'étude "Maison Dégradée" pour le mois de janvier

Les caractéristiques de la dispersion pour le mois de janvier sont dans le tableau 4.9, la maison dégradée a un besoin de chauffage, sur les deux zones, supérieur à 70% aux besoins de chauffage observés pour le modèle de la maison DM original. On pouvait s'attendre à une incertitude relative moindre pour un besoin plus important, or la dispersion a elle aussi augmenté avec les besoins et l'incertitude relative est du même ordre de grandeur pour les deux configurations.

**Table 4.9 – Modèle Dégradé en janvier** Valeurs de dispersion du besoin de chauffage total et par zone

<b>Résultats Maison Dégradée</b>				
<b>Chauffage</b>	<b>Moyenne <math>\mu</math></b>	<b>Ecart type <math>\sigma</math></b>	<b>Intervalle de confiance</b>	<b>Variation</b>
RDJ [ $.10^8$ Joule]	13,5	1,35	$13,5 \pm 2,6$	19,6%
Etage [ $.10^8$ Joule]	14,3	1,34	$14,3 \pm 2,6$	18,4%
Totale [ $kWh.m^{-2}$ ]	7,91	0,761	$7,91 \pm 1,49$	18,9%
<b>Récapitulatif cas test original</b>				
<b>Chauffage</b>	<b>Moyenne <math>\mu</math></b>	<b>Ecart type <math>\sigma</math></b>	<b>Intervalle de confiance</b>	<b>Variation</b>
RDJ [ $.10^8$ Joule]	7,96	0,990	$7,96 \pm 1,94$	24,4%
Etage [ $.10^8$ Joule]	8,36	0,841	$8,36 \pm 1,65$	19,7%
Totale [ $kWh.m^{-2}$ ]	4,65	0,516	$4,65 \pm 1,01$	21,7%

On va étudier les causes de cette dispersion, pour cela on estime les indices de sensibilité qui sont présentés sous forme d'histogramme en figure 4.24. On observe une augmentation de l'influence de la température sur la dispersion des besoins qui est significative pour le RDJ mais moindre pour l'Etage.



**Figure 4.24 – Maison Dégradée en janvier** Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance calculées par 100 répliques bootstrap. On observe une influence du rayonnement direct plus importante au rez-de-jardin qu'à l'étage. La température est essentiellement responsable des variations de besoin de chauffage.

Les valeurs des effets principaux  $S_i$  ainsi que l'estimé bootstrap  $\hat{S}_i$  sont regroupés dans le tableau 4.10.

**Table 4.10 – Maison Dégradée en janvier** Valeurs des  $S_i$  et  $\hat{S}_i$  pour le besoin de chauffage total et de chaque zone

	Résultats Maison Dégradée					
	RDJ		Etage		Total	
	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$
$H_R$	0,003	-0,026	-0,001	-0,024	0,001	-0,024
$T$	<b>0,825</b>	<b>0,820</b>	<b>0,949</b>	<b>0,953</b>	<b>0,895</b>	<b>0,895</b>
$R_{Dir}$	<b>0,179</b>	<b>0,153</b>	<b>0,068</b>	<b>0,049</b>	<b>0,116</b>	<b>0,087</b>
$R_{Diff}$	0,041	0,013	0,026	0,006	0,033	0,014
$V_{Dir}$	0,027	-0,001	0,021	0,002	0,025	0,002
$V_{Vit}$	0,015	-0,015	-0,009	-0,027	0,002	-0,018
	Récapitulatif cas test original					
	RDJ		Etage		Total	
	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$
$H_R$	0,008	-0,028	0,003	-0,029	0,006	-0,028
$T$	<b>0,680</b>	<b>0,667</b>	<b>0,905</b>	<b>0,892</b>	<b>0,796</b>	<b>0,790</b>
$R_{Dir}$	<b>0,301</b>	<b>0,274</b>	<b>0,103</b>	<b>0,081</b>	<b>0,199</b>	<b>0,168</b>
$R_{Diff}$	0,053	0,019	0,030	0,017	0,042	0,011
$V_{Dir}$	0,018	-0,013	0,0164	-0,006	0,019	-0,006
$V_{Vit}$	0,029	0,000	-0,001	-0,017	0,015	-0,003

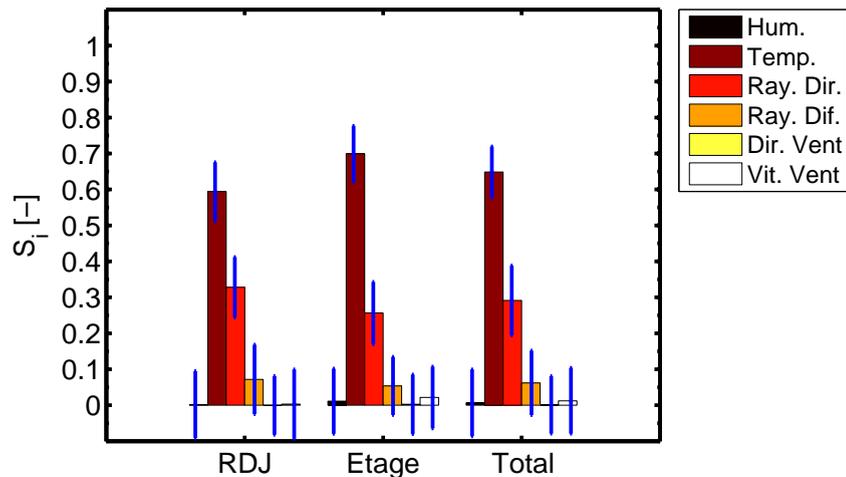
#### 4.3.2.3 Résultats de l'étude "Maison Dégradée" pour le mois de juillet

Les dispersions sur le besoin de refroidissement sont évaluées pour le mois de juillet et regroupées dans le tableau 4.11.

Selon les résultats du tableau 4.11, les besoins de refroidissement du modèle dégradé augmentent à l'Etage de 28% par rapport au cas test original mais seulement de 8% au RDJ. Contrairement au cas du Triple Vitrage, beaucoup de composants ont été modifiés dans ce modèle dégradé et il devient plus difficile d'interpréter la disparité entre zones. La différence de l'augmentation des besoins entre les zones peut être due aux connections avec le vide sanitaire ou bien avec les combles. L'incertitude relative est du même ordre de grandeur. On estime les indices de sensibilité, les résultats sont sous forme d'histogramme en figure 4.25.

**Table 4.11 – Maison Dégradée en juillet** Valeur de dispersion du besoin de chauffage total et par zone

	Résultats Maison Dégradée			
Refroidissement	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $\cdot 10^8$ Joule]	6,00	0,866	$6,00 \pm 1,70$	28,3%
Etage [ $\cdot 10^8$ Joule]	5,44	0,871	$5,44 \pm 1,71$	31,4%
Totale [ $kWh.m^{-2}$ ]	3,26	0,494	$3,26 \pm 0,97$	29,7%
	Récapitulatif cas test original			
Refroidissement	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $\cdot 10^8$ Joule]	5,57	0,711	$5,57 \pm 1,39$	25%
Etage [ $\cdot 10^8$ Joule]	4,24	0,640	$4,24 \pm 1,25$	29,6%
Totale [ $kWh.m^{-2}$ ]	2,79	0,384	$2,79 \pm 0,75$	27,0%



**Figure 4.25 – Maison Dégradée en juillet** Indice de sensibilité principal sur le besoin de chauffage des deux zones et le besoin total. Bar d'erreur de  $1.96\sigma$  correspondant à 95% de seuil de confiance calculé par 100 répliques bootstrap. On observe une influence du rayonnement direct plus importante au rez-de-jardin qu'à l'étage et la température extérieure est essentiellement responsable des variations de besoins de refroidissement.

Les valeurs des effets principaux  $S_i$  ainsi que  $\hat{S}_i$  sont regroupées dans le tableau 4.12. On observe une augmentation de la sensibilité de la température qui est du même ordre de grandeur pour l'étage et le rez-de-jardin contrairement au cas hivernal.

**Table 4.12 – Maison Dégradée en juillet** Valeurs des indices de sensibilité  $S_i$  et de son estimé bootstrap  $\hat{S}_i$  pour le besoin de chauffage total et de chaque zone

	Résultats Maison Dégradée					
	RDJ		Etage		Total	
	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$
$H_R$	0,002	0,002	0,011	0,015	0,006	0,007
$T$	<b>0,594</b>	<b>0,605</b>	<b>0,699</b>	<b>0,691</b>	<b>0,648</b>	<b>0,654</b>
$R_{Dir}$	<b>0,328</b>	<b>0,333</b>	<b>0,256</b>	<b>0,258</b>	<b>0,291</b>	<b>0,293</b>
$R_{Diff}$	0,072	0,073	0,054	0,061	0,062	0,070
$V_{Dir}$	-0,001	0,001	0,003	0,004	0,001	0,001
$V_{Vit}$	0,003	0,014	0,021	0,032	0,012	0,016
	Récapitulatif cas test original					
	RDJ		Etage		Total	
	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$	$S_i$	$\hat{S}_i$
$H_R$	-0,023	-0,009	-0,015	-0,006	-0,019	-0,007
$T$	<b>0,483</b>	<b>0,490</b>	<b>0,582</b>	<b>0,591</b>	<b>0,530</b>	<b>0,530</b>
$R_{Dir}$	<b>0,391</b>	<b>0,404</b>	<b>0,323</b>	<b>0,334</b>	<b>0,359</b>	<b>0,377</b>
$R_{Diff}$	0,062	0,070	0,038	0,045	0,049	0,054
$V_{Dir}$	-0,007	0,018	-0,004	0,022	-0,006	0,020
$V_{Vit}$	-0,021	-0,002	-0,010	0,003	-0,016	0,002

Nous avons montré l'évolution de l'incertitude associée à la variabilité météorologique pour différents types de bâtiment et l'influence de la température augmente en fonction des cas. Dans le cas de la Maison Dégradée cette augmentation s'explique par la diminution de l'isolation avec l'extérieur mais dans le cas "Triple Vitrage" il est moins intuitif de comprendre cette diminution. Nous pensons que l'effet "filtre solaire" du triple vitrage diminue l'effet du rayonnement solaire ce qui induit une augmentation de l'effet de la température. Afin de visualiser l'évolution du rayonnement direct en fonction de la température extérieure nous étudions l'évolution des indices sur une sortie temporelle.

### 4.3.3 Comparaison de la variabilité temporelle des indices : effet du stockage sur les indices de sensibilité

Nous voulons visualiser l'évolution des indices de sensibilité en fonction du temps pour les différents modèles. L'évolution temporelle apporte des informations supplémentaires sur l'amplitude de variation des indices et leur simultanéité. On calcul l'influence de chaque variable météo sur les besoins horaires. On compare pour les deux saisons le cas original, le cas "Triple Vitrage" et le modèle de maison DM dégradé.

Nous avons vu dans les analyses précédentes que la température et le rayonnement direct sont les variables responsables de la dispersion de sortie. Le comportement nocturne et diurne de chaque indice n'est pas visible à partir des valeurs des indices de sensibilité sur une sortie agrégée mensuelle. La présence du rayonnement solaire diminue l'effet de la température. On observe ce comportement en opposition de phase sur l'ensemble des graphes d'évolution des indices en particulier sur la période de janvier.

Sur les graphes concernant le cas test original en figure 4.26 et le cas "Triple Vitrage" en janvier en figure 4.28, on observe un régime transitoire au début du signal qui est moins important pour le cas "Maison Dégradée" (figure 4.26), il ne dure que trois jours comparé aux sept jours des autres signaux. Nous pensons que cette transition est due à l'initialisation des simulations et que l'écart entre Maison Dégradée et le cas test original est causé par la différence de capacité de stockage du bâtiment : l'inertie. Avant la phase de simulation et donc de calcul des performances, EnergyPlus répète le premier jour de simulation jusqu'à ce que le bâtiment se stabilise. L'initialisation va dépendre de la convergence et du nombre de jours d'initialisation fixés. Dans le cas de l'étude, seule la convergence est fixée afin d'initialiser au mieux les simulations et le logiciel répète la première journée tant qu'il n'obtient pas la convergence. Or, nous pensons que cette procédure surestime la variabilité de la première journée et à cause de l'inertie, elle influence la sensibilité des jours suivants. Lorsque l'inertie est moins importante dans le cas "Maison Dégradée" (figure 4.27 et figure 4.30) l'influence de l'initialisation est moins importante.

Un autre témoin de la différence de stockage de l'énergie entre le cas test original et le cas "Maison Dégradée" est l'influence du rayonnement solaire durant la nuit. Ce phénomène est particulièrement visible sur le cas test original au mois de juillet en figure 4.29 car la valeur de la sensibilité du rayonnement direct durant la nuit est au minimum d'environ 0,3. Dans le cas où la quantité de rayonnement entrant est limitée par le triple vitrage, cette valeur diminue et est proche de 0,2 (figure 4.31). Dans le cas "Maison Dégradée" moins inertielle les valeurs sont plus basses et atteignent des valeurs nulles (figure 4.30) alors que l'amplitude des variations de sensibilité diurne et nocturne augmente de 0,2 par rapport au cas inertiel.

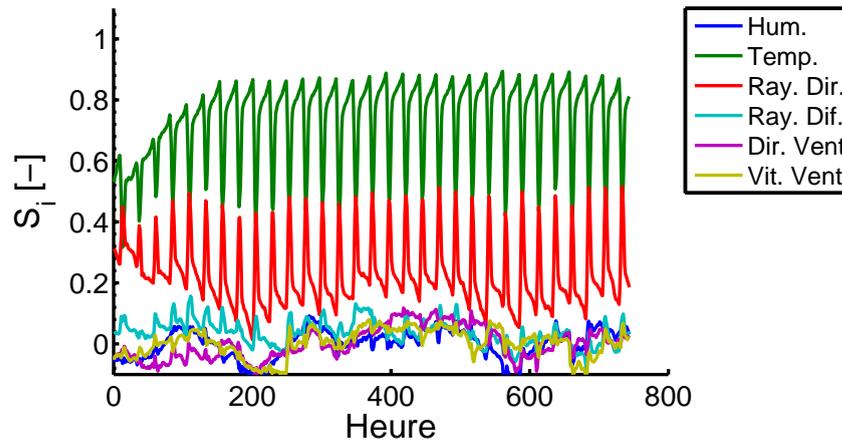


Figure 4.26 – Cas test original en janvier Évolution des indices de sensibilité en fonction du besoin de chauffage total horaire

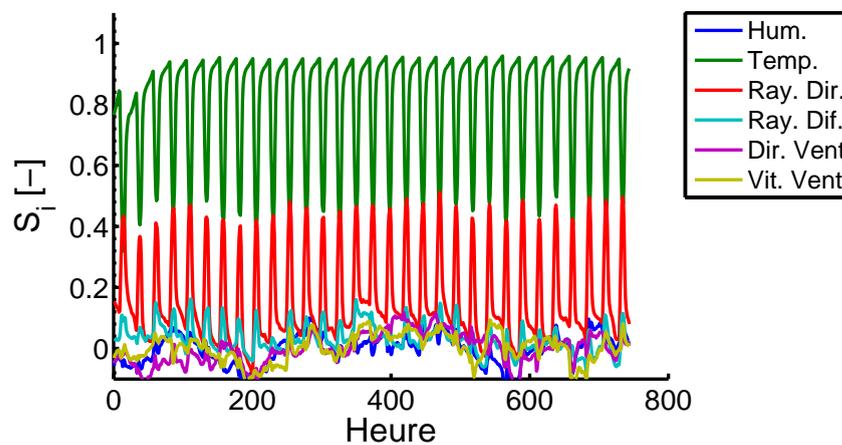


Figure 4.27 – Maison Dégradée en janvier Évolution des indices de sensibilité en fonction du besoin de chauffage total horaire

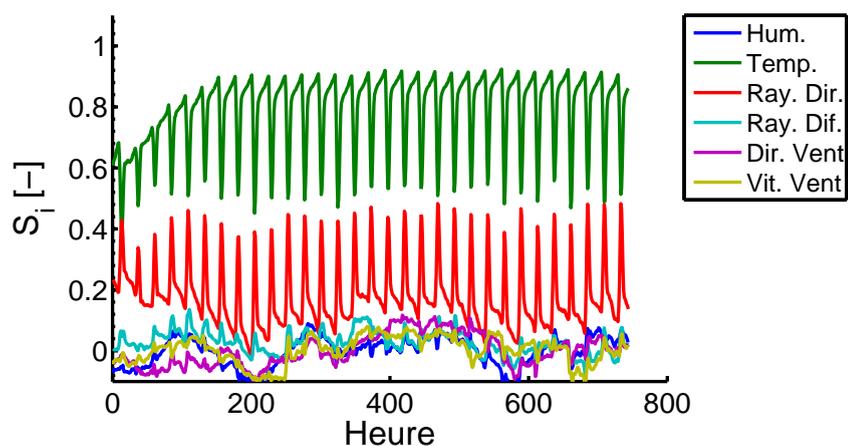


Figure 4.28 – Triple Vitrage en janvier Évolution des indices de sensibilité en fonction du besoin de chauffage total horaire

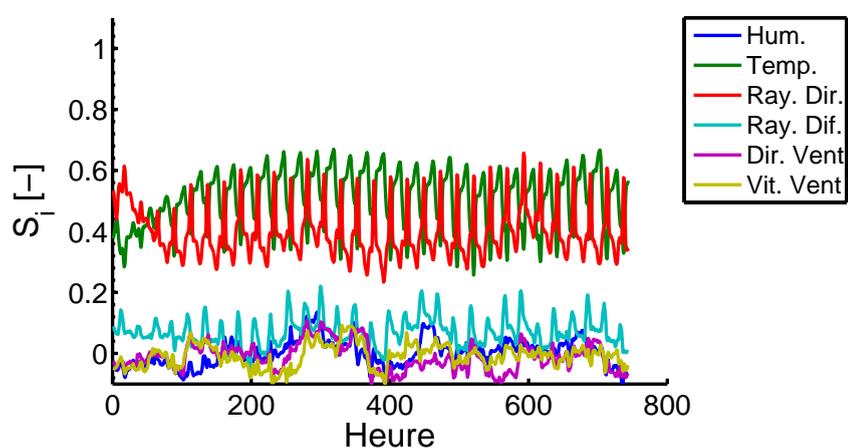
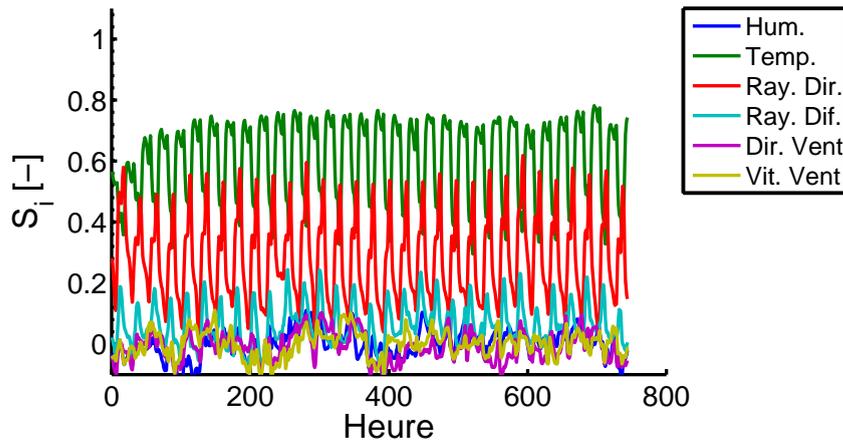
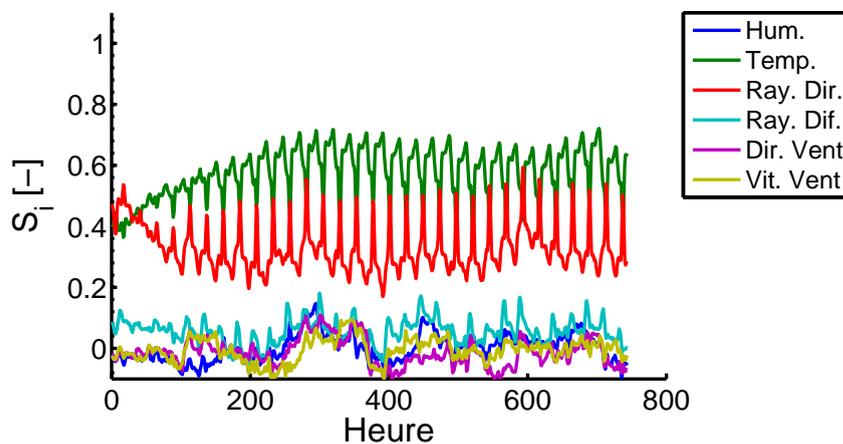


Figure 4.29 – Cas test original en juillet Évolution des indices de sensibilité en fonction du besoin de refroidissement total horaire



**Figure 4.30 – Maison dégradée en juillet** Évolution des indices de sensibilité en fonction du besoin de refroidissement total horaire. On observe la différence d'amplitude et de valeur de sensibilité pour le rayonnement direct la nuit par rapport au cas test original lié à la différence d'inertie entre les deux modèles



**Figure 4.31 – Triple Vitrage en juillet** Évolution des indices de sensibilité en fonction du besoin de refroidissement total horaire

Les résultats de dispersion et de paramètres influents dépendent du modèle que l'on étudie. A travers différents cas de modifications du modèle thermique, d'un seul composant (Triple Vitrage) à plusieurs (Maison Dégradée), nous illustrons l'impact de ces modifications sur la dispersion et les indices de sensibilité. Les dispersions que l'on observe sont conséquentes et causées en été et hiver en majorité par la température. Qu'en est-il dans le cas d'incertitude associée au modèle thermique. Quelle dispersion et quelle sensibilité allons-nous obtenir en considérant les incertitudes des entrées statiques ?

## 4.4 Variabilité et sensibilité du modèle pour des entrées statiques et temporelles

On cherche à évaluer les différentes variabilités de sortie lorsque des paramètres statiques et dynamiques sont en jeu dans l'analyse de sensibilité. Pour cela nous allons effectuer une analyse de sensibilité en ne prenant en compte que les paramètres statiques puis une analyse avec les deux types d'entrées : statiques et temporelles. Le climat utilisé est celui de Lyon et le modèle thermique est le cas test original. En simulation thermique dynamique, il n'est pas rare qu'un modèle de bâtiment comporte plus d'une centaine de paramètres. Il s'agit des paramètres géométriques, relatifs au système, à l'occupation, aux scénarios, aux paramètres thermo-physiques, à l'environnement, etc. Pour limiter le nombre de paramètres dans l'analyse de sensibilité, on en sélectionne 49 représentatifs des caractéristiques des vitrages, de l'isolation, de la structure et de l'environnement. Ces 49 paramètres, leur valeurs nominales ainsi que la fonction de distribution associée sont présentés en annexe.

Les 49 paramètres sont échantillonnés pour la majorité sur une loi normale de moyenne, leur valeur nominale  $x^*$  et d'écart type  $\sigma$  tel que  $x^* = \bar{x}$ ;  $x = \bar{x} \pm 10\%$  à un niveau de confiance 95% soit  $1,96\sigma$ .  $\mathcal{N}[\bar{x}, \frac{0,1\bar{x}}{2}]$ . Pour l'infiltration et l'albédo une loi uniforme est choisie, les incertitudes des infiltrations sont de 10%  $\mathcal{U}[x^* - 0.1x^*, x^* + 0.1x^*]$ . L'albédo varie de 0,2 à 0,4 soit un écart de l'ordre de 33% par rapport à la valeur nominale 0,3.

La méthode d'échantillonnage LHS est utilisée pour générer une matrice  $\mathbf{X}^{(0)}$  de  $N \times p$  avec  $N = 1000$  et  $p = 49$  de laquelle est obtenue la seconde matrice  $\mathbf{X}^{(1)}$  par permutation des  $X_i^{(0)}$  voir section 2.2.3.1. Les deux matrices d'échantillons sont propagées dans le modèle pour obtenir après  $2N$  évaluations du modèle  $Y_k^{(0)}$  et  $Y_k^{(1)}$  avec  $k$  la  $k$ -ième sortie. On souhaite tout d'abord évaluer la dispersion et la sensibilité en ne prenant en compte que la variabilité des entrées statiques.

### 4.4.1 Analyses de sensibilité et d'incertitude pour les entrées statiques

#### 4.4.1.1 Résultats de l'étude statique pour le mois de janvier

Le tableau 4.13 regroupe les caractéristiques de la dispersion pour le besoin de chauffage. Les valeurs moyennes sont légèrement plus importantes (3-4%) que les valeurs moyennes du cas test original avec variabilité météo. La valeur de l'écart type est trois fois moins importante, l'incertitude relative est donc faible, de l'ordre de 7%. A partir de la dispersion, même faible, nous évaluons les paramètres les plus influents parmi les 49. Un grand nombre de paramètres a une valeur de sensibilité comprise entre 0,03 et 0,05. Sachant que l'écart type sur la variabilité de l'estimateur  $S_i$ , tout comme les études précédentes est de l'ordre de 0,05, nous considérons un effet principal significatif lorsqu'il est supérieur à 0,05. Le tableau 4.14 renseigne l'ordre d'influence, le nom de la variable et sa valeur de sensibilité  $S_i$ .

**Table 4.13 – Statique en janvier** Valeurs de dispersion du besoin de chauffage total et par zone

Chauffage	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $.10^8$ Joule]	8,33	0,326	$8,33 \pm 0,64$	7,67%
Etage [ $.10^8$ Joule]	8,63	0,277	$8,63 \pm 0,54$	6,29%
Totale [ $kWh.m^{-2}$ ]	4,83	0,160	$4,93 \pm 0,31$	6,49%

Pour la zone RDJ, les deux premiers paramètres influents concernent les caractéristiques des vitrage sud : la conductance (*cond.sud* : 21%) et le facteur solaire (*fact.sol.sud* : 19%). Puis le reste de la dispersion est expliqué à 18% par la conductivité de l'isolation des murs extérieurs (*cond.isol.murext.*), à 15% par l'albédo et enfin à 9% par les infiltrations du RDJ. La somme des indices de sensibilité est de 0,82, cette perte d'informations peut être due au fait que de nombreux paramètres sont faiblement influents et contiennent cependant de l'information. Il est difficile d'évaluer si cette influence est réelle ou due à la variabilité de l'estimateur. Cependant, il est aussi probable que le manque d'information soit dû à des interactions notamment entre le facteur solaire Sud et l'albédo.

Les paramètres influents sur la dispersion du besoin de chauffage à l'Etage sont principalement la conductivité de l'isolation des murs extérieurs (53%) et les infiltrations de l'Etage (20%). A part l'albédo (7%) les paramètres influents sont tous en rapport avec les déperditions thermiques contrairement au rez-de-jardin (facteur solaire Sud et albédo) et on retrouve cette sensibilité plus accrue du RDJ au solaire par rapport à l'Etage.

**Table 4.14 – Statique en janvier** Variables les plus influentes et valeurs des indices de sensibilité  $S_i$  pour le besoin de chauffage total et de chaque zone

Rang	RDJ		Etage		Total	
	Variable	$S_i$	Variable	$S_i$	Variable	$S_i$
1	<i>fact.sol.sud</i>	0,209	<i>cond.isol.murext.</i>	0,525	<i>cond.isol.murext.</i>	0,360
2	<i>cond.sud</i>	0,194	<i>infilt.etage</i>	0,196	<i>cond.sud</i>	0,141
3	<i>cond.isol.murext.</i>	0,180	<i>cond.ouest</i>	0,073	<i>fact.sol.sud</i>	0,125
4	<i>albédo</i>	0,149	<i>cond.sud</i>	0,067	<i>albédo</i>	0,124
5	<i>infil.rdj</i>	0,088	<i>albédo</i>	0,065	<i>cond.ouest</i>	0,055
6					<i>infilt.etage</i>	0,051

#### 4.4.1.2 Résultats de l'étude statique pour le mois de juillet

Le tableau 4.15 regroupe les valeurs de dispersion sur les besoins de refroidissement pour le mois de juillet. Les valeurs moyennes sont proches des valeurs du cas test original, l'écart type augmente ce qui a pour effet d'augmenter l'incertitude relative qui reste cependant proche des 30%. L'identification des paramètres responsables de la dispersion révèle que l'albédo domine à plus de 85% les besoins sur les deux zones (tableau 4.16). L'albédo est la variable d'entrée qui est la plus dispersée avec 33% de variation et sa forte influence en été montre une nouvelle fois la sensibilité du bâtiment au rayonnement solaire. Au regard des résultats précédent notamment sur le stockage de chaleur (influence du rayonnement solaire la nuit), l'albédo influence la quantité de rayonnement réfléchi et donc la sensibilité au stockage. Quelle influence de l'albédo allons nous observer si la variabilité des données météorologiques sont également prise en compte dans l'analyse ? La variabilité sur le rayonnement solaire et son influence serait-elle plus importante que celle de l'albédo ? Est-ce la quantité de rayonnement initial qui influence la réponse du bâtiment ou la part de rayonnement réfléchi ?

**Table 4.15 – Statique en juillet** Valeur de dispersion du besoin de refroidissement total et par zone

Refroidissement	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $.10^8$ Joule]	5,77	0,972	$5,77 \pm 1,91$	33%
Etage [ $.10^8$ Joule]	4,39	0,612	$4,39 \pm 1,20$	27,3%
Totale [ $kWh.m^{-2}$ ]	2,89	0,451	$2,89 \pm 0,88$	30,6%

**Table 4.16 – Statique en juillet** Variables les plus influentes et valeurs des indices de sensibilité  $S_i$  pour le besoin de refroidissement total et de chaque zone

Rang	RDJ		Etage		Total	
	Variable	$S_i$	Variable	$S_i$	Variable	$S_i$
1	<i>albédo</i>	0,858	<i>albédo</i>	0,878	<i>albédo</i>	0,868
2	<i>fact.sol.sud</i>	0,097	<i>fact.sol.sud</i>	0,072	<i>fact.sol.sud</i>	0,087

#### 4.4.2 Analyses de sensibilité et d'incertitude avec des entrées temporelles et statiques

##### 4.4.2.1 Résultats de l'étude mixte des entrées statiques et temporelles pour le mois de janvier

L'ajout de la variabilité statique à la variabilité météorologique n'induit pas de différence significative et on obtient les mêmes valeurs moyennes tandis que la dispersion augmente légèrement ce qui induit une incertitude relative un peu plus importante mais toujours dans l'ordre de grandeur des 20%-25%. On visualise, dans le tableau 4.18 le peu d'impact de l'introduction de la variabilité statique : les variables influentes et leur valeur de sensibilité sont les mêmes que pour l'étude temporelle seule. La conductivité de l'isolation apparaît toutefois dans les paramètres influents. La variabilité météorologique domine les dispersions de sortie.

**Table 4.17 – Statique et dynamique en janvier** Valeurs de dispersion du besoin de chauffage total et par zone

Chauffage	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $.10^8$ Joule]	7,96	1,06	$7,96 \pm 2,08$	26,1%
Etage [ $.10^8$ Joule]	8,36	0,887	$8,36 \pm 1,74$	20,8%
Totale [ $kWh.m^{-2}$ ]	4,65	0,547	$4,65 \pm 1,07$	23,1%

**Table 4.18 – Statique et Dynamique en janvier** Variables les plus influentes et valeurs des indices de sensibilité  $S_i$  pour le besoin de chauffage total et de chaque zone

Rang	RDJ		Etage		Total	
	Variable	$S_i$	Variable	$S_i$	Variable	$S_i$
1	$T$	0,605	$T$	0,806	$T$	0,709
2	$R_{Dir}$	0,259	$R_{Dir}$	0,084	$R_{Dir}$	0,168
3			<i>cond.isol.murext.</i>	0,071	<i>cond.isol.murext.</i>	0,051

4.4.2.2 Résultats de l'étude mixte des entrées statiques et temporelles pour le mois de juillet

Les valeurs moyennes sont, comme pour le mois de janvier, égales aux valeurs du cas test original avec entrées temporelles par contre, la dispersion observée dans le tableau 4.19 est importante, l'incertitude relative est donc conséquente. Nous déterminons les raisons de cette augmentation en estimant les indices de sensibilité (tableau 4.20) afin de comparer les entrées influentes à celles obtenues dans le cas de l'étude des entrées temporelles seules. L'albédo explique la moitié des variations du besoin de refroidissement total et son influence est plus importante pour le RDJ que pour l'Etage. Puis les variables influentes sont les variables météorologiques rayonnement direct ( $R_{Dir}$ ) et la température extérieure ( $T$ ) plus ou moins influentes l'une par rapport à l'autre en fonction de la zone. Ainsi la variabilité météorologique est influente sur la réponse d'un modèle thermique mais les entrées statiques qui modulent leur importance au sein du bâtiment peuvent également être très influentes.

Table 4.19 – Statique et Dynamique en juillet Valeurs de dispersion du besoin de refroidissement total et par zone

Refroidissement	Moyenne $\mu$	Ecart type $\sigma$	Intervalle de confiance	Variation
RDJ [ $\cdot 10^8$ Joule]	5,58	1,21	$5,58 \pm 2,37$	42,5%
Etage [ $\cdot 10^8$ Joule]	4,24	0,893	$4,24 \pm 1,75$	41,3%
Totale [ $kWh.m^{-2}$ ]	2,80	0,597	$2,80 \pm 1,17$	41,8%

Table 4.20 – Statique et Dynamique en juillet Variables les plus influentes et valeurs des indices de sensibilité  $S_i$  pour le besoin de refroidissement total et de chaque zone

Rang	RDJ		Etage		Total	
	Variable	$S_i$	Variable	$S_i$	Variable	$S_i$
1	<i>albédo</i>	0,545	<i>albédo</i>	0,402	<i>albédo</i>	0,488
2	$R_{Dir}$	0,161	$T$	0,266	$T$	0,190
3	$T$	0,139	$R_{Dir}$	0,187	$R_{Dir}$	0,173
4	<i>fact.sol.sud</i>	0,077			<i>fact.sol.sud</i>	0,064

## 4.5 Conclusion et tableau des dispersions

Nous avons exploré la dispersion et la sensibilité des entrées météorologiques dans différents cas. Nous avons montré que la valeur de dispersion des variables météorologiques ainsi que la valeur moyenne ou la zone d'exploration impactent sur les résultats de dispersion de sortie et de sensibilité. Nous avons pu apprécier la modification de réponse du bâtiment entre Lyon et Strasbourg et visualiser l'impact de la température plus variable et plus froide de Strasbourg par rapport à Lyon.

A partir de cet échantillon du climat lyonnais nous avons confronté les résultats de dispersion et de sensibilité de différents modèles thermiques. Nous avons vu que le stockage du bâtiment joue un rôle dans les valeurs de sensibilité obtenues alors que la sensibilité du rayonnement solaire dépend de l'inertie du bâtiment. Lors d'une étude avec prise en compte de la variabilité de 49 entrées statiques nous avons démontré que la variabilité des entrées statiques peuvent induire des modifications importantes dans la détermination de la sensibilité des entrées temporelles. Le tableau 4.21 regroupe l'ensemble des valeurs de dispersion pour les besoins de chauffage et de refroidissement totaux pour les différents cas ainsi que les entrées influentes et leur valeur de sensibilité, avec en gras les valeurs du cas test original qui sert de comparaison tout le long de ce chapitre.

Table 4.21 – Récapitulatif des dispersions pour les différents cas

<b>CHAUFFAGE</b>			
	<b>Dispersion sur le besoin total</b>		<b>Sensibilité</b>
Cas	Int. de conf. (95%)	Inc. rel. (95%)	Paramètre influent( $S_i$ )
<b>Lyon</b>	4,65 ± 1,01	<b>21,1%</b>	<i>T</i> ( <b>0,80</b> ) <i>R<sub>Dir</sub></i> ( <b>0,20</b> )
<b>Strasbourg</b>	4,98 ± 1,60	32,1%	<i>T</i> (0,87) <i>R<sub>Dir</sub></i> (0,09)
<b>Triple Vitrage</b>	4,42 ± 0,89	20,1%	<i>T</i> (0,84) <i>R<sub>Dir</sub></i> (0,15)
<b>Maison Dégradée</b>	7,91 ± 1,49	18,9%	<i>T</i> (0,90) <i>R<sub>Dir</sub></i> (0,12)
<b>Etude Statique</b>	4,83 ± 0,31	6,49%	<i>cond.isol.murext.</i> (0,36) <i>cond.sud</i> (0,14)...
<b>Etude Mixte</b>	4,65 ± 1,07	23,1%	<i>T</i> (0,71) <i>R<sub>Dir</sub></i> (0,17)
<b>REFROIDISSEMENT</b>			
	<b>Dispersion sur le besoin total</b>		<b>Sensibilité</b>
Cas	Int. de conf. (95%)	Inc. rel. (95%)	Paramètre influent( $S_i$ )
<b>Lyon</b>	2,79 ± 0,75	<b>27,0%</b>	<i>T</i> ( <b>0,53</b> ) <i>R<sub>Dir</sub></i> ( <b>0,36</b> )
<b>Strasbourg</b>	1,57 ± 0,78	49,2%	<i>R<sub>Dir</sub></i> (0,59) <i>T</i> (0,31)
<b>Triple Vitrage</b>	1,84 ± 0,66	36,1%	<i>T</i> (0,59) <i>R<sub>Dir</sub></i> (0,31) <i>R<sub>Diff</sub></i> (0,05)
<b>Maison Dégradée</b>	3,26 ± 0,97	29,7%	<i>T</i> (0,65) <i>R<sub>Dir</sub></i> (0,29)
<b>Etude Statique</b>	2,89 ± 0,88	30,6%	<i>alb.</i> (0,87) <i>fact.sol.sud</i> (0,09)
<b>Etude Mixte</b>	2,80 ± 1,17	41,8%	<i>alb.</i> (0,49) <i>T</i> (0,19) <i>R<sub>Dir</sub></i> (0,17)



# Conclusion

Les sollicitations telles que les occupants, la météo ou encore les scénarios de consommation des usages font partie des entrées les plus incertaines et potentiellement les plus influentes sur les performances d'un bâtiment basse consommation. Il est nécessaire pour pouvoir garantir des performances de déterminer les dispersions de sortie associées à la variabilité des entrées temporelles et d'en déterminer les variables responsables pour mieux réduire leur variabilité ou encore concevoir le bâtiment de manière robuste.

Pour répondre à cette problématique, nous avons utilisé les indices de sensibilité de Sobol qui font partie des méthodes d'analyse de sensibilité adaptées aux modèles complexes à grandes dimensions tels que les modèles de bâtiment pour la simulation thermique dynamique. La gestion des entrées fonctionnelles étant un verrou scientifique pour les méthodes d'analyse de sensibilité standard, une méthodologie originale a été développée dans le cadre de cette thèse afin de générer des échantillons compatibles avec l'estimation de la sensibilité. Bien que la méthode soit générique aux entrées fonctionnelles, elle a été validée dans ce travail de thèse pour le cas des données météorologiques et tout particulièrement à partir des fichiers météo moyens utilisés en simulation thermique dynamique.

Les deux aspects principaux de ce travail de développement résident dans la caractérisation de la variabilité des données météorologiques et dans la génération des échantillons permettant l'estimation de la sensibilité de chaque variable météorologique sur la dispersion des performances d'un bâtiment.

Nous avons montré que les échantillons sont cohérents avec l'hypothèse d'indépendance dans la génération des entrées temporelles corrélées, ce qui permet d'appliquer les techniques d'analyse de sensibilité. Par ailleurs, nous obtenons des échantillons représentatifs de la variabilité naturelle des données météorologiques sur dix ans. Nous observons cependant quelques anomalies dans la génération liées à la perte d'informations sur les structures corrélatives, qui restent toutefois minimales. Une étude comparative avec des données météorologiques mesurées sur le long terme permettra d'évaluer avec précision les lacunes de la méthodologie.

La propagation des échantillons nous permet d'évaluer la dispersion sur les besoins énergétiques de la maison INCAS-DM. Quel que soit le climat ou le type de modèle thermique testé, la température et le rayonnement direct dominant la dispersion en été et hiver. L'ordre de grandeur de ces dispersions est de 20% à 30% par rapport à la valeur moyenne en considérant un intervalle de confiance à 95%. Une étude faisant intervenir l'incertitude des entrées statiques et dynamiques du bâtiment montre même une dispersion en été de plus de 40% due majoritairement à l'influence de l'albédo. Par ailleurs nous avons montré que l'étude de l'évolution temporelle des indices de sensibilité des variables météorologiques permettait d'évaluer la capacité de stockage du bâtiment en mettant en valeur la sensibilité du rayonnement direct sur la dispersion des besoins en période nocturne.

Les résultats présentés ici sont issus du cas particulier de la maison INCAS-DM pour les climats de Lyon et de Strasbourg mais la méthodologie est bien sûr applicable à tous bâtiments et climats tempérés afin d'évaluer la dispersion sur les besoins énergétiques liée à la variabilité météo et d'en déterminer les variables responsables. Nous avons vu qu'elle pouvait être combinée à une analyse de sensibilité comprenant également les entrées statiques telles que les paramètres de conception. Par ailleurs cette méthodologie peut être adaptée à d'autres entrées fonctionnelles corrélées comme les profils d'occupation, les scénarios pour permettre une étude globale des incertitudes et de la sensibilité en thermique du bâtiment et bien entendu également dans un cadre différent de celui de la thermique du bâtiment.

Dans ces travaux, l'étude des besoins énergétiques a permis plus de transparence sur les capacités de l'enveloppe ce qui est une première étape dans l'évaluation des performances énergétiques.

# Perspectives

Les perspectives de ce travail nous permettront d'envisager des diagnostics de performance prenant en compte les incertitudes ce qui est une condition indispensable à la mise en place de la garantie de performance.

En effet, la garantie de performance passe par le calcul des consommations et non pas uniquement des besoins ce qui sous-entend connaître la gestion du bâtiment et les systèmes énergétiques en tenant compte de leurs incertitudes. Ainsi il sera nécessaire d'étudier les performances du bâtiment en prenant en compte son utilisation réelle avec le profil d'occupation, les scénarios (ouverture/fermeture de volet, température de consigne adaptées, etc.), l'efficacité des systèmes, le renouvellement d'air, etc.

Il faudra également être capable d'estimer la dispersion sur les consommations liées à la variabilité des scénarios, de l'occupation, des systèmes, de la météo et des paramètres statiques, ce qui permettra d'une part d'obtenir un intervalle de confiance sur les résultats et d'autre part de recenser les paramètres les plus influents. Cette approche pourra mener à des choix de systèmes énergétiques plus stables face aux variations induites par l'occupation ou les conditions météorologiques.

Pour permettre ce type d'étude des travaux sur l'utilisation des modèles réduits devront être menés afin de réduire le temps de calcul associé à l'analyse de sensibilité pour être capable de gérer des bâtiments complexes.



# Bibliographie

- [1] F Anstett-Collin, T Mara, L Denis-Vidal, and J Goffart. UASA of complex models : Coping with dynamic and static inputs. In *7th International Conference on Sensitivity Analysis of Model Output , July 1-4 2013 , Nice*, 2013.
- [2] L. Bertolo and B. Bourges. Données climatiques utilisées dans le bâtiment. *Les Techniques de l'Ingénieur*, 1992.
- [3] Mahabir Bhandari, Som Shrestha, and Joshua New. Evaluation of weather datasets for building energy simulation. *Energy and Buildings*, 49 :109–118, 2012.
- [4] D. P. Bloomfield. An investigation into analytical and empirical validation techniques for dynamic thermal models of buildings. *Internal report of BRE*, 1988.
- [5] H. Breesch and a. Janssens. Performance evaluation of passive cooling in office buildings based on uncertainty and sensitivity analysis. *Solar Energy*, 84(8) :1453–1467, August 2010.
- [6] A.L.S. Chan. Developing a modified typical meteorological year weather file for hong kong taking into account the urban heat island effect. *Building and Environment*, 46(12) :2434–2441, 2011.
- [7] G. Chastaing, F. Gamboa, and C. Prieur. Generalized hoeffding-sobol decomposition for dependent variables. application to sensitivity analysis. *Electronic Journal of Statistics*, 6 :2420–2448, 2012.
- [8] Lou Chesné, Thierry Duforestel, Jean-jacques Roux, and Gilles Rusaouën. Energy saving and environmental resources potentials : Toward new methods of building design. *Building and Environment*, 58 :199–207, 2012.
- [9] J A Clarke, P P Yaneske, and A A Pinney. The Harmonisation of Thermal Properties of Building Materials. *BRE Publication BEPAC Research Report*, 2 :1–87, 1990.
- [10] R. Confalonieri. Combining a weather generator and a standard sensitivity analysis method to quantify the relevance of weather variables on agrometeorological models inputs. *Theoretical and Applied Climatology*, 108 :19–30, 1996.

- 
- [11] Drury B Crawley. Which Weather Data Should You Use for Energy Simulations of Commercial Buildings? *ASHRAE Transactions*, 104 Part 2 :1–18, 1998.
- [12] Drury B. Crawley, Linda K. Lawrie, Frederick C. Winkelmann, W.F. Buhl, Y. Joe Huang, Curtis O. Pedersen, Richard K. Strand, Richard J. Liesen, Daniel E. Fisher, Michael J. Witte, and Jason Glazer. EnergyPlus : creating a new-generation building energy simulation program. *Energy and Buildings*, 33(4) :319–331, April 2001.
- [13] L.W. Crow. Weather year for energy calculations. *ASHRAE Journal*, 6, 1984.
- [14] Pieter de Wilde and Wei Tian. Management of thermal performance risks in buildings subject to climate change. *Building and Environment*, 55(0) :167 – 177, 2012.
- [15] M. Eames, T. Kershaw, and D. Coley. The appropriate spatial resolution of future weather files for building simulation. *Journal of Building Performance Simulation*, 5(6) :347–358, 2012.
- [16] B. Efron and R. Tibshirani. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1 :54–75, 1986.
- [17] J.M. Filkenstein and R.E. Schafer. Improved goodness to fit tests. *Journal Biometrica*, 58, 1971.
- [18] Jeanne Goffart, Etienne Wurtz, Gérard Sauce, and Timea Bejat. Impact and Source of Uncertainties in High Efficiency Building Simulation : Some Examples. In *Twelfth International IBPSA Conference Proceedings*, pages 2822–2827, Sydney, Australia, 2011.
- [19] A. Guaus, A. Bsaibes, T. Cartailleur, Prieur C., E. Lebon, and Frederic G. Time-dependent sensitivity and uncertainty analyses of an agroclimatic model for the water status management of vineyard. In *7th International Conference on Sensitivity Analysis of Model Output , July 1-4 2013 , Nice*, 2013.
- [20] Gilles Guyon. Role of the model user in results obtained from simulation software program. In *Building Simulation Conference, Prague, Czech Republic*, 1997.
- [21] Frédéric Haldi and Darren Robinson. The impact of occupants’ behaviour on building energy demand. *Journal of Building Performance Simulation*, 4(4) :323–338, 2011.
- [22] J C Helton and F J Davis. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, 81 :23–69, 2003.

- 
- [23] T. Homma and A. Saltelli. Importance measures in global sensitivity analysis of model output. *Reliability Engineering & System Safety*, 52(1) :1–17, 1996.
- [24] Tianzhen Hong, S.K Chou, and T.Y Bong. Building simulation : an overview of developments and information sources. *Building and Environment*, 35(4) :347–361, May 2000.
- [25] Christina J. Hopfe and Jan L.M. Hensen. Uncertainty analysis in building performance simulation for design support. *Energy and Buildings*, 43(10) :2798–2805, July 2011.
- [26] S.P. Huang, S.T. Quek, and K.K. Phoon. Convergence study of the truncated karhunen-loeve expansion for simulation of stochastic processes. *International Journal for Numerical Methods in Engineering*, 52(9) :1029–1043, 2001.
- [27] R. Iman and W. Conover. A distribution free approach to inducing a rank correlation among input variables. *Communications in Statistics - Simulation and Computation*, 11(3) :311–334, 1982.
- [28] Jensen, S. Ø. Validation of building energy simulation programs : a methodology. *Energy and Buildings*, 22(2) :133 – 144, 1995.
- [29] Mark F. Jentsch, Patrick a.B. James, Leonidas Bourikas, and AbuBakr S. Bahaj. Transforming existing weather data for worldwide locations to enable energy and building performance simulation under future climates. *Renewable Energy*, 55 :514–524, July 2013.
- [30] Mark F. Jentsch, AbuBakr S. Bahaj, and Patrick a.B. James. Climate change future proofing of buildings. Generation and assessment of building simulation weather files. *Energy and Buildings*, 40(12) :2148–2168, January 2008.
- [31] R.D. Judkoff. Validation of building energy analysis simulation programs at the solar energy research institute. *Energy and Buildings*, 10(3) :221 – 239, 1988.
- [32] M. Kolokotroni, X. Ren, M. Davies, and A. Mavrogianni. London’s urban heat island : Impact on current and future energy consumption in office buildings. *Energy and Buildings*, 47 :302–311, 2012.
- [33] Jens-Peter Kreiss and Efstathios Paparoditis. Bootstrap methods for dependent data : A review. *Journal of the Korean Statistical Society*, 40(4) :357–378, December 2011.

- [34] S. Kucherenko, S. Tarantola, and P. Annoni. Estimation of global sensitivity indices for models with dependent variables. *Computer Physics Communications*, 183(4) :937–946, 2012.
- [35] H. Li, J. Harvey, and A. Kendall. Field measurement of albedo for different land cover materials and effects on thermal performance. *Building and Environment*, 59(0) :536 – 546, 2013.
- [36] L. Lilburne and S. Tarantola. Sensitivity analysis of spatial models. *Int. J. of Geographical Information Science*, 23(2) :151–168, February 2009.
- [37] Iain Macdonald and Paul Strachan. Practical application of uncertainty analysis. *Energy and Buildings*, 33 :0–8, 2001.
- [38] Thierry Mara. Développement et Applications d’Outils d’Aide à la Vérification, la Calibration, l’Analyse de Sensibilité et d’Incertitude de Modèles. Technical report, Université de la réunion, 2008.
- [39] Thierry Mara and Onimihamina Rakoto Joseph. Comparison of some efficient methods to evaluate the main effect of computer model factors. *Journal of Statistical Computation and Simulation*, 78(2) :167–178, February 2008.
- [40] Météo-France. <http://www.climat.meteofrance.com>.
- [41] Eugene a. Mohareb, Christopher a. Kennedy, L.D. Danny Harvey, and Kim D. Pressnail. Decoupling of building energy use and climate. *Energy and Buildings*, 43(10) :2961–2963, October 2011.
- [42] Vahid M. Nik and Angela Sasic Kalagasidis. Impact study of the climate change on the energy performance of the building stock in stockholm considering four climate uncertainties. *Building and Environment*, 60 :291–304, 2013.
- [43] Vahid M. Nik, Angela Sasic Kalagasidis, and Erik Kjellstrom. Statistical methods for assessing and analysing the building performance in respect to the future climate. *Building and Environment*, 53 :107–118, 2012.
- [44] U.S Department of Energy. <http://apps1.eere.enregy.gov/buildings/energyplus/>.
- [45] U.S Department of Energy. [http://apps1.eere.enregy.gov/buildings/energyplus/weatherdata\\_about.cfm](http://apps1.eere.enregy.gov/buildings/energyplus/weatherdata_about.cfm).
- [46] Elena Palomo, Del Barrio, and Gilles Guyon. Theoretical basis for empirical model validation using parameters space analysis tools. *Energy and Buildings*, 35 :985–996, 2003.

- 
- [47] Elena Palomo and Gilles Guyon. Application of parameters space analysis tools for empirical model validation. *Energy and Buildings*, 36 :23–33, 2004.
- [48] Wout Parys, Dirk Saelens, and Hugo Hens. Coupling of dynamic building simulation with stochastic modelling of occupant behaviour in offices- a review-based integrated methodology. *Journal of Building Performance Simulation*, 4(4) :339–358, 2011.
- [49] S. Patidar, D.P. Jenkins, G.J. Gibson, and P.F.G. Banfill. Analysis of probabilistic climate projections : heat wave, overheating and adaptation. *Journal of Building Performance Simulation*, 6(1) :65–77, 2013.
- [50] K K Phoon, S P Huang, and S T Quek. Implementation of Karhunen - Loeve expansion for simulation using a wavelet-Galerkin scheme. *Probabilistic Engineering Mechanics*, 55 :293–303, 2002.
- [51] Konstantin Protassov. *Analyse statistique des données expérimentales*. EDP Sciences, 2008.
- [52] Mickael Rabouille, Etienne Wurtz, and Pascal Perrotin. Analysis of dynamic thermal simulation for refurbishment. In *Thirteenth International IBPSA Conference Proceedings*,, Chambéry, France, 2013.
- [53] Olivier Roustant. *Produits dérivés climatiques : aspects économétriques et financiers*. PhD thesis, Université Claude Bernard Lyon 1, 2003.
- [54] Nathalie Saint-geours and Christian Lavergne. Analyse de sensibilité globale d’un modèle spatialisé pour l’évaluation économique du risque d’inondation. *Journal de la Société Française de Statistique*, 152(1) :49–71, 2011.
- [55] A. Saltelli, M. Ratto, S. Tarantola, and F. Campolongo. Update 1 of : Sensitivity analysis for chemical models. *Chemicals Reviews*, 112 :1–21, 2012.
- [56] Andrea Saltelli and Paola Annoni. How to avoid a perfunctory sensitivity analysis. *Environmental Modelling and Software*, 25(12) :1508–1517, 2010.
- [57] Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. Variance based sensitivity analysis of model output . Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2) :259–270, 2010.
- [58] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. Introduction to Sensitivity Analysis. In *Global Sensitivity Analysis. The Primer*, pages 1–51. John Wiley & Sons, Ltd, 2008.

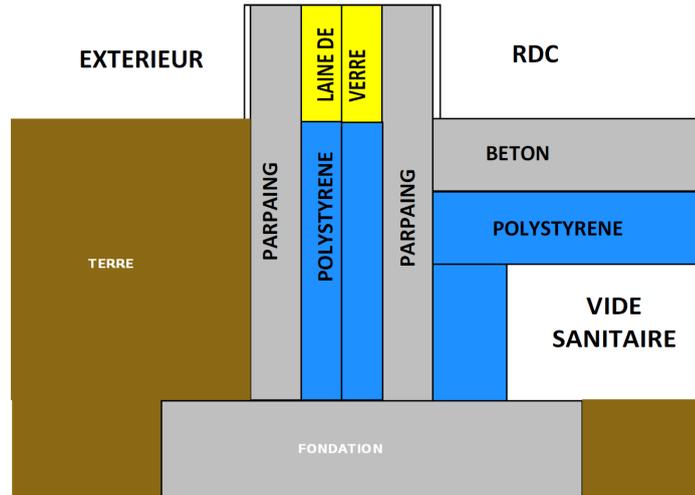
- 
- [59] I Sobol, S Tarantola, D Gatelli, S Kucherenko, and W Mauntz. Estimating the approximation error when fixing unessential factors in global sensitivity analysis. *Reliability Engineering & System Safety*, 92(7) :957–960, July 2007.
- [60] I. M. Sobol'. Sensitivity estimates for nonlinear mathematical models. *Math. Mod. and Comput. Exp.*, 1 :407–414, 1993.
- [61] Clara Spitz. *Analyse de la fiabilité des outils de simulation et des incertitudes de métrologie appliquée à l'efficacité énergétique des bâtiments*. PhD thesis, Université de Grenoble, 2012.
- [62] Clara Spitz, Laurent Mora, Etienne Wurtz, and Arnaud Jay. Practical application of uncertainty analysis and sensitivity analysis on an experimental house. *Energy and Buildings*, 55 :459–470, December 2012.
- [63] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and System Safety*, 93 :964–979, 2008.
- [64] Jun Tanimoto, Aya Hagishima, Takeshi Iwai, and Naoki Ikegaya. Total utility demand prediction for multi-dwelling sites by a bottom-up approach considering variations of inhabitants' behaviour schedules. *Journal of Building Performance Simulation*, 6(1) :53–64, 2013.
- [65] Wei Tian. A review of sensitivity analysis methods in building energy analysis. *Renewable and Sustainable Energy Reviews*, 20(2013) :411–419, April 2013.
- [66] Wei Tian and Pieter de Wilde. Uncertainty and sensitivity analysis of building performance using probabilistic climate projections : A uk case study. *Automation in Construction*, 20(8) :1096–1109, 2011.
- [67] Marija Trčka and Jan L.M. Hensen. Overview of HVAC system simulation. *Automation in Construction*, 19(2) :93–99, March 2010.
- [68] Liping Wang, Paul Mathew, and Xiufeng Pang. Uncertainties in energy consumption introduced by building operations and weather for a medium-size office building. *Energy and Buildings*, 53 :152–158, 2012.

# Annexes

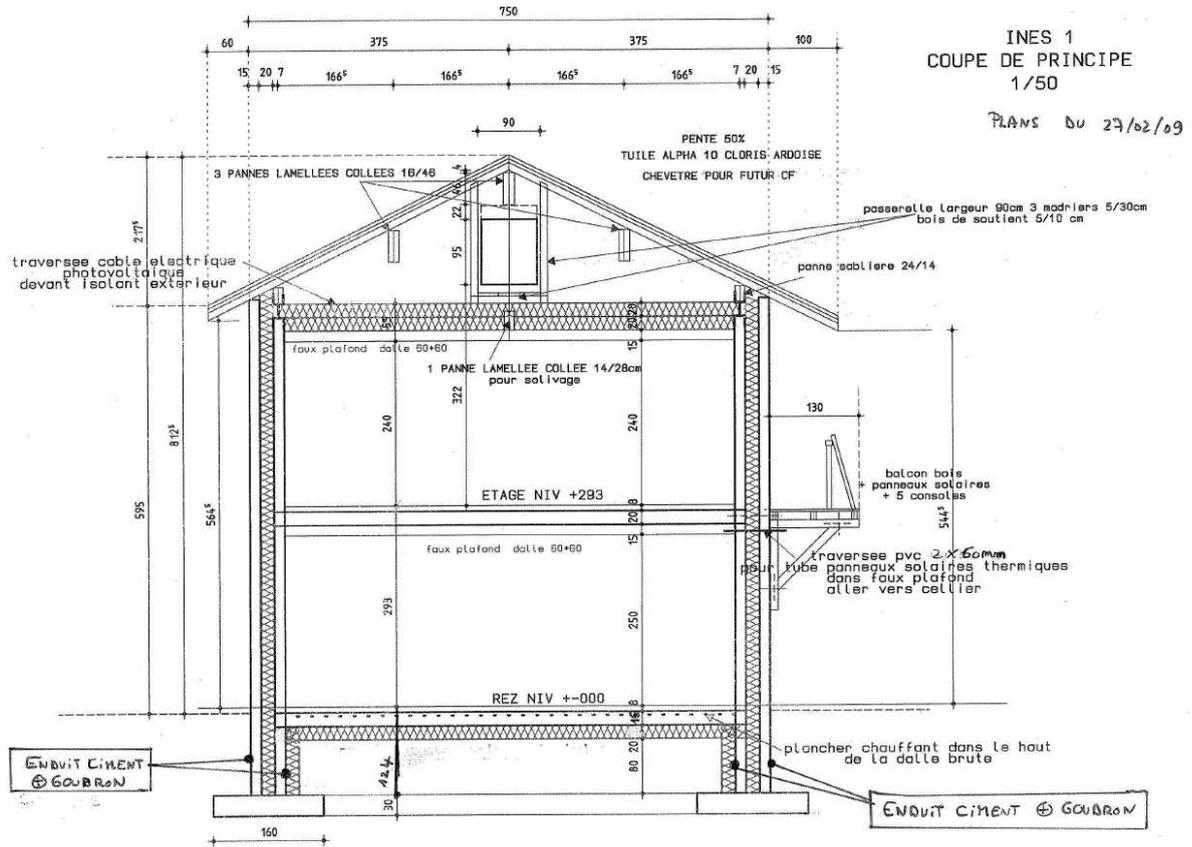
## Plan et composition de la maison INCAS

Composition des parois et propriétés thermo-physiques des matériaux de la maison INCAS-DM

Type de paroi	Matériau	e [cm]	$\lambda$ [W.m <sup>-1</sup> .K <sup>-1</sup> ]	$\rho$ [kg.m <sup>-3</sup> ]	C <sub>p</sub> [J.kg <sup>-1</sup> .K <sup>-1</sup> ]
Mur extérieur	Enduit extérieur	1	1	1450	500
	Parpaing	15	0,74	800	648
	Laine de verre	20	0,035	12	840
	Parpaing	15	0,74	800	648
	Plâtre	1,5	0,42	1200	837
Plancher bas	Enduit extérieur	1	1	1450	500
	Polystyrène extrudé	20	0,030	35	1400
	Dalle pleine	16	2,5	2400	1000
	Chape béton	8	1,75	2400	880
Combles	Laine de verre	40	0,035	12	840
	Plâtre	1,5	0,42	1200	837
Plancher intermédiaire	Poutrelle	16	1,23	1300	648
	Dalle pleine	4	2,5	2400	1000
	Chape béton	8	1,75	2400	880
	Carrelage	1	0,41	1200	1000



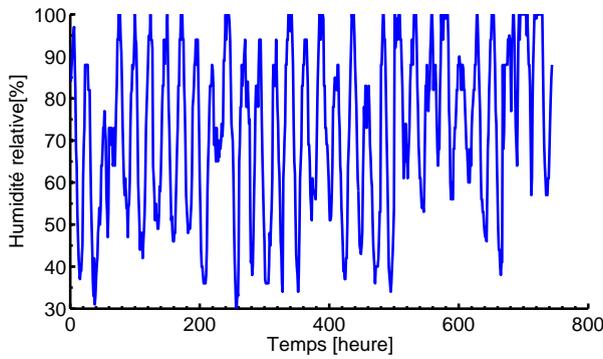
Liaison plancher bas mur extérieur de la maison I-DM [61]



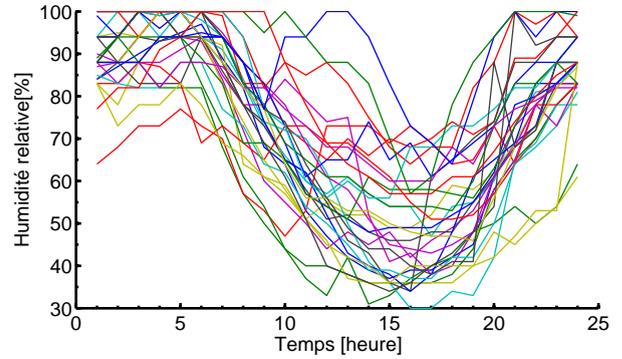
Coupe de principe de la maison I-DM

## Données statistiques extraites du fichier IWEC Lyon au mois de juillet

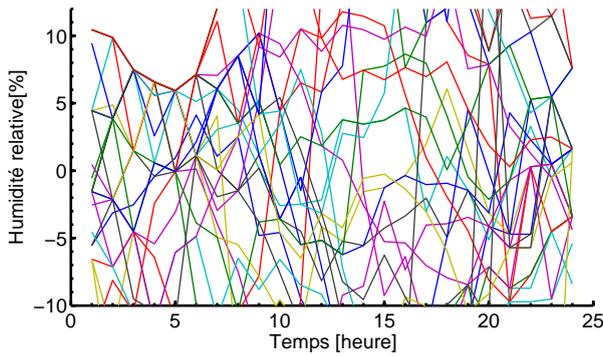
- Humidité relative
- Rayonnement Direct
- Rayonnement Diffus
- Vitesse de vent
- Direction de vent



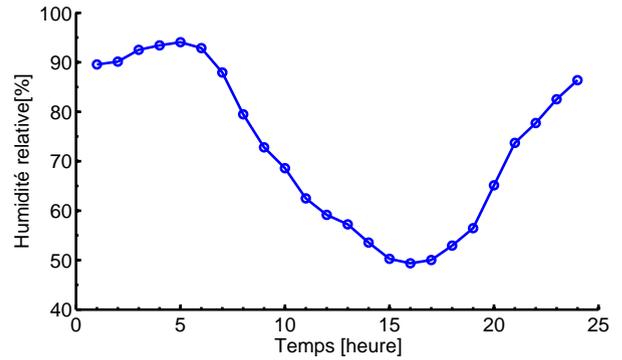
(a) Signal source



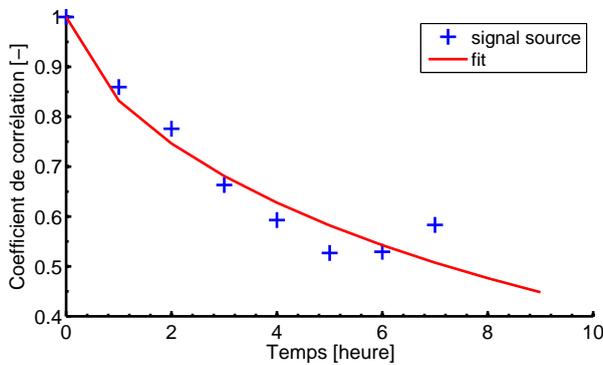
(b) Signal concaténé



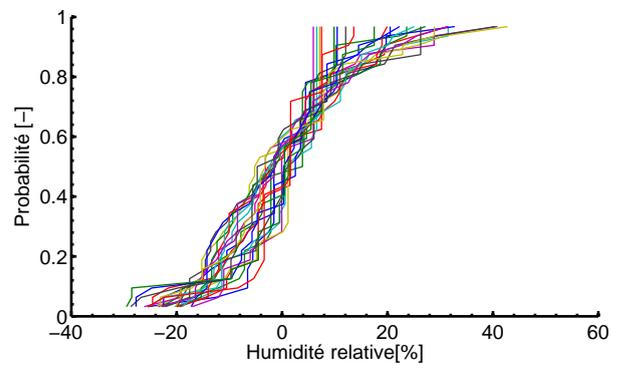
(c) Résidu



(d) Moyenne horaire

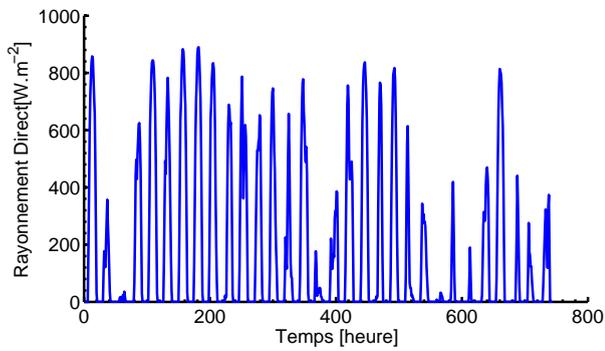


(e) Fonction d'auto-corrélation

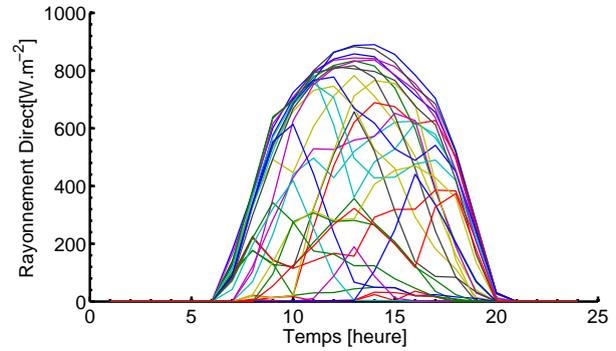


(f) Fonctions de distribution

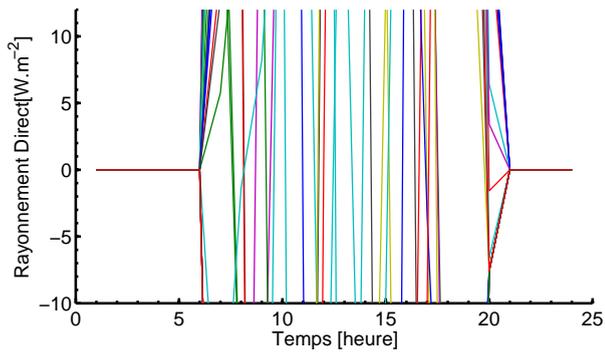
Données statistiques extraites pour l'humidité relative



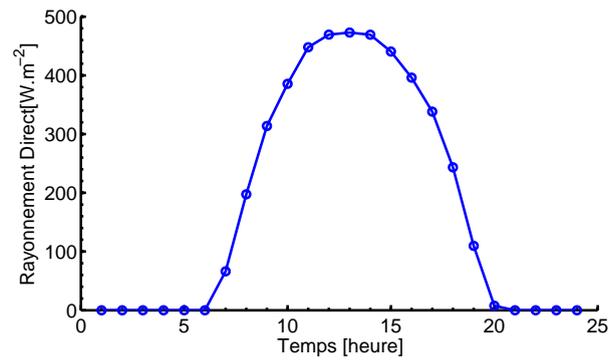
(a) Signal source



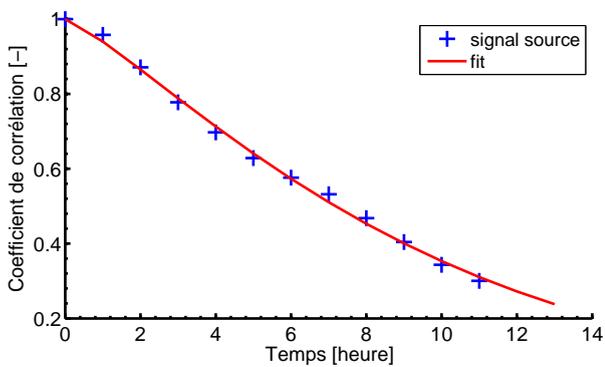
(b) Signal concaténé



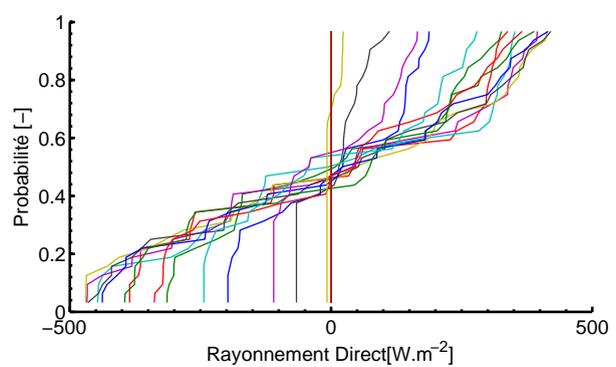
(c) Résidu



(d) Moyenne horaire

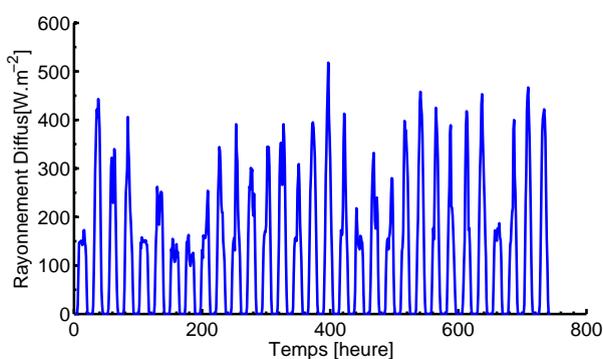


(e) Fonction d'auto-corrélation

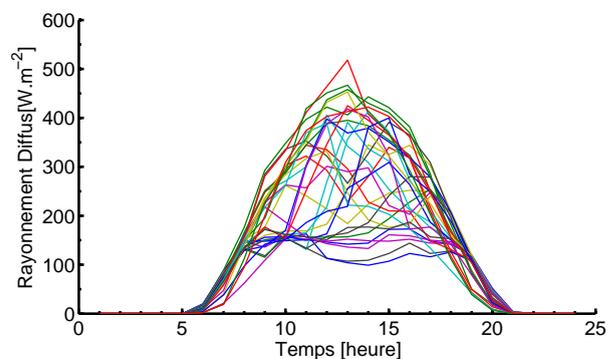


(f) Fonctions de distribution

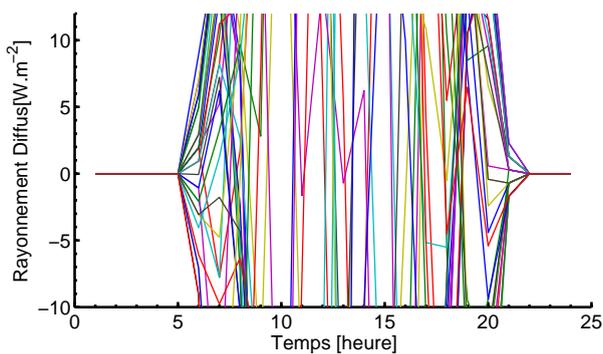
Données statistiques extraites pour le **rayonnement direct**



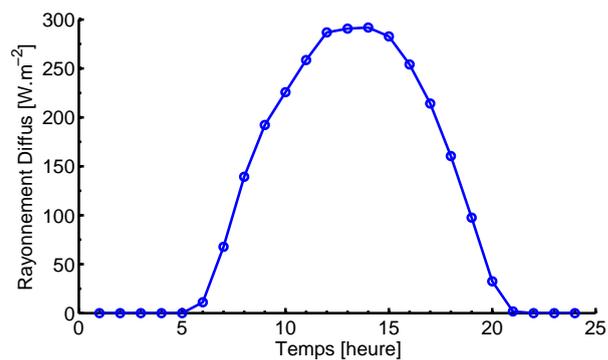
(a) Signal source



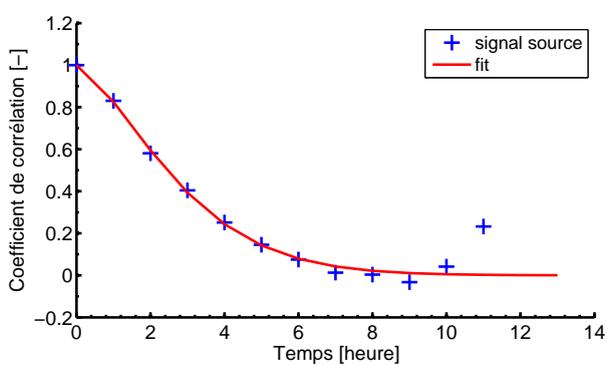
(b) Signal concaténé



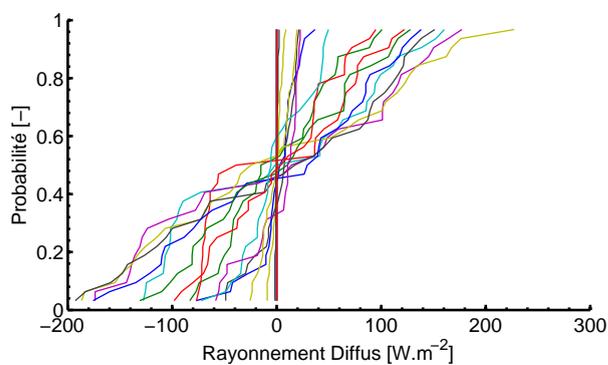
(c) Résidu



(d) Moyenne horaire

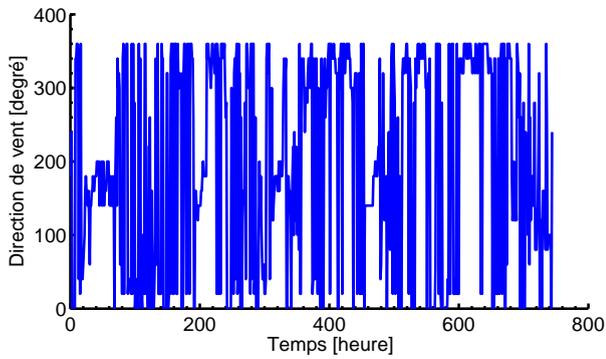


(e) Fonction d'auto-corrélation

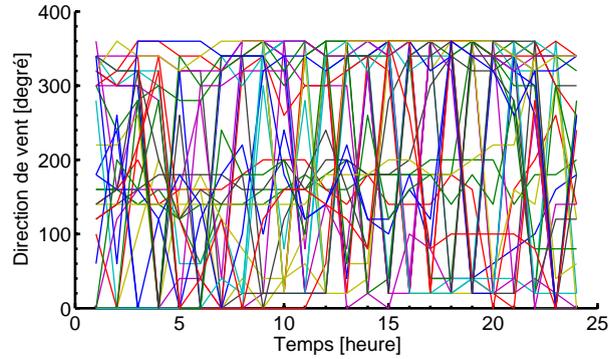


(f) Fonctions de distribution

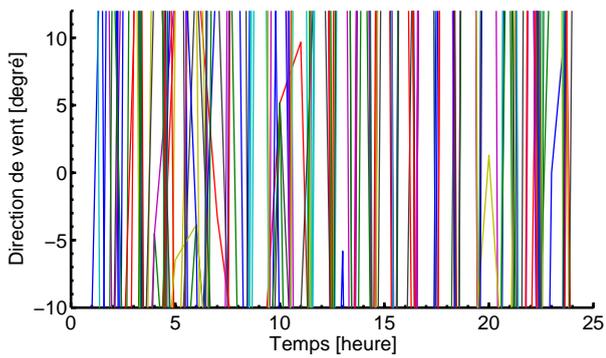
Données statistiques extraites pour le rayonnement diffus



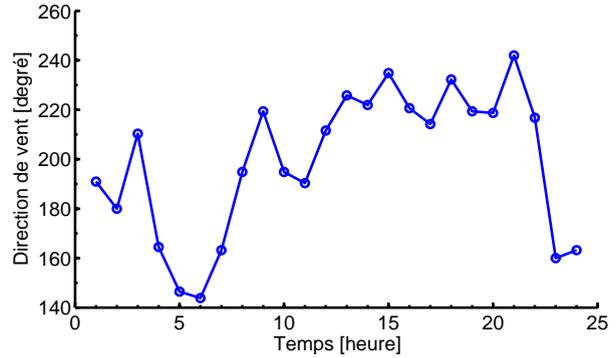
(a) Signal source



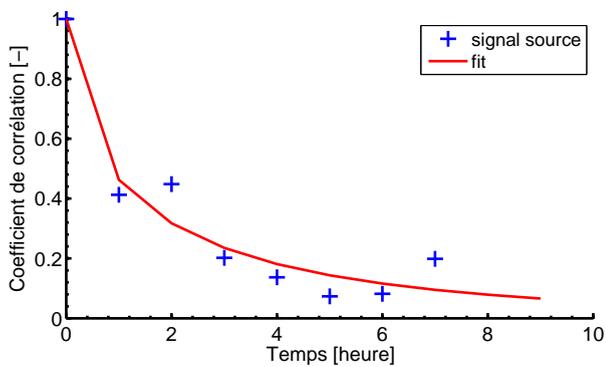
(b) Signal concaténé



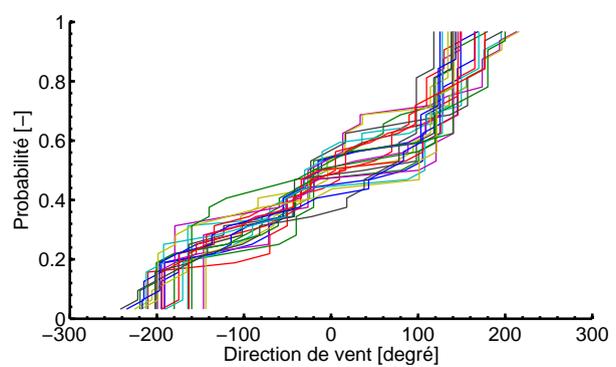
(c) Résidu



(d) Moyenne horaire

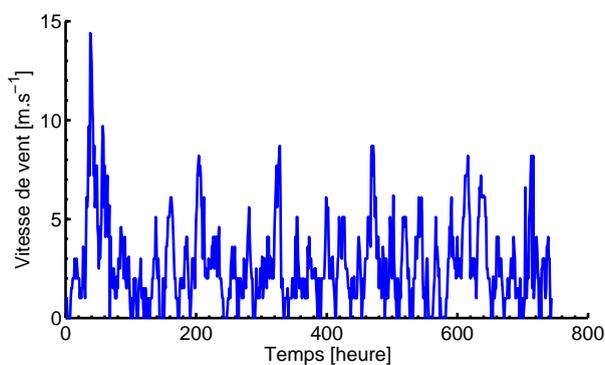


(e) Fonction d'auto-corrélation

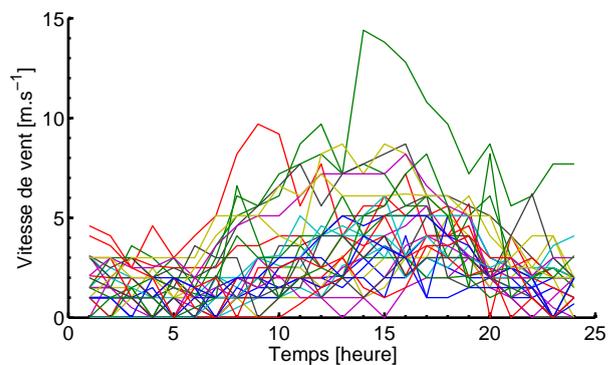


(f) Fonctions de distribution

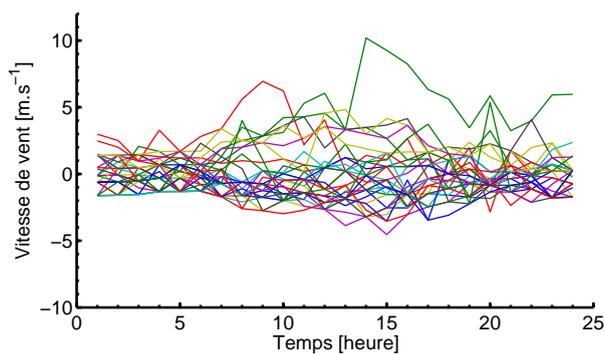
Données statistiques extraites pour la **direction de vent**



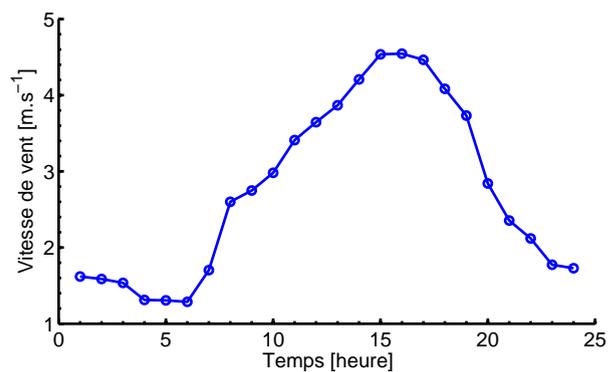
(a) Signal source



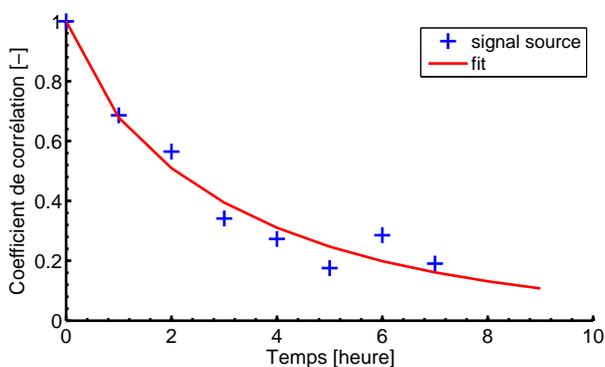
(b) Signal concaténé



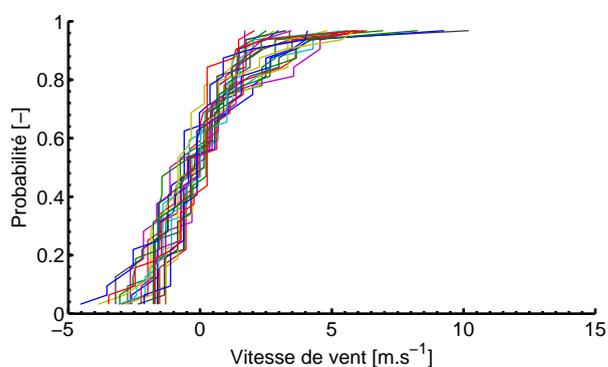
(c) Résidu



(d) Moyenne horaire



(e) Fonction d'auto-corrélation



(f) Fonctions de distribution

Données statistiques extraites pour la vitesse de vent

## Valeurs et dispersions des 49 paramètres statiques de l'analyse de sensibilité statique et mixte

Valeurs et dispersions des 49 paramètres statiques de l'analyse de sensibilité statique et mixte

Composant	Paramètre	abréviation	valeur
laine de verre 35 isolation murs ext.	conductivité masse volumique	cond.isol.murext rho.isol.murext	$\mathcal{N}(0.035, 0.00175)$ $\mathcal{N}(12, 0.6)$
laine de verre 32 Isolation mur comble	conductivité masse volumique	cond.isol.murcomb rho.isol.murcomb	$\mathcal{N}(0.032, 0.0016)$ $\mathcal{N}(12, 0.6)$
laine de verre 35 Isolation Etage-comble	conductivité masse volumique	cond.isol.Et.comble rho.isol.murcomb	$\mathcal{N}(0.035, 0.00175)$ $\mathcal{N}(12, 0.6)$
Polystyrene extrudé Isolant VS-RDJ	conductivité masse volumique	cond.isol.VS.rdj rho.isol.VS.rdj	$\mathcal{N}(0.03, 0.0015)$ $\mathcal{N}(35, 1.75)$
Roofmate Isolant murs VS	conductivité masse volumique	cond.isol.mur.vs rho.isol.mur.vs	$\mathcal{N}(0.029, 0.00145)$ $\mathcal{N}(35, 1.75)$
Parpaing murs	conductivité Masse volumique	cond.parpaing rho.parpaing	$\mathcal{N}(1, 0.05)$ $\mathcal{N}(1100, 55)$
Chape Plancher bas	conductivité Masse volumique	cond.chape.bas rho.chape.bas	$\mathcal{N}(2.75, 0.1375)$ $\mathcal{N}(2400, 120)$
Dalle 16 cm Plancher bas	conductivité Masse volumique	cond.dalle.bas16 rho.dall.bas	$\mathcal{N}(2.5, 0.125)$ $\mathcal{N}(2400, 120)$
Poutrelle Plancher haut	conductivité Masse volumique	cond.poutre rho.poutre	$\mathcal{N}(1.23, 0.0615)$ $\mathcal{N}(1300, 65)$
Chape Plancher haut	conductivité Masse volumique	cond.chape.haut rho.chape.haut	$\mathcal{N}(2.75, 0.1375)$ $\mathcal{N}(2400, 120)$
Dalle 4 cm Plancher haut	conductivité Masse volumique	cond.dalle.haut4 rho.dalle.haut4	$\mathcal{N}(2.5, 0.125)$ $\mathcal{N}(2400, 120)$

Composant	Paramètre	abréviation	valeur
carrelage Sol RDC	conductivité Masse volumique	cond.carrelage Rho.carrelage	$\mathcal{N}(0.41, 0.0205)$ $\mathcal{N}(1200, 60)$
Enduit Mur ext.	conductivité Masse volumique	cond.enduit rho.enduit	$\mathcal{N}(1, 0.05)$ $\mathcal{N}(1450, 72.5)$
Plâtre Mur int.	conductivité masse volumique	cond.platre rho.platre	$\mathcal{N}(0.43, 0.0215)$ $\mathcal{N}(1200, 60)$
Faux plafond Plafond Etage plafond RDJ	conductivité Masse volumique	cond.fxplaf rho.fxplaf	$\mathcal{N}(0.35, 0.0175)$ $\mathcal{N}(36.9, 1.845)$
Tuile toit	conductivité Masse volumique	Cond.tuile Rho.tuile	$\mathcal{N}(2.2, 0.11)$ $\mathcal{N}(2700, 135)$
Air plafond (fx plafond)	conductivité Masse volumique	Cond.air Rho.air	$\mathcal{N}(0.025, 0.00125)$ $\mathcal{N}(1.23, 0.0615)$
gravier Sol environnant	conductivité Masse volumique	Cond.gravier Rho.gravier	$\mathcal{N}(2, 0.1)$ $\mathcal{N}(1600, 80)$
Fenêtre sud	Conductance transmittance	cond.sud fact.sol.sud	$\mathcal{N}(1.1, 0.055)$ $\mathcal{N}(0.6, 0.03)$
Fenêtre est	Conductance transmittance	cond.est fact.sol.est	$\mathcal{N}(1.1, 0.055)$ $\mathcal{N}(0.6, 0.03)$
Fenêtre ouest	Conductance transmittance	cond.ouest fact.sol.ouest	$\mathcal{N}(1.1, 0.055)$ $\mathcal{N}(0.6, 0.03)$
Fenêtre nord (triple vitrage)	conductance transmittance	cond.nord fact.sol.nord	$\mathcal{N}(0.7, 0.035)$ $\mathcal{N}(0.45, 0.0225)$
Albédo		Albédo	$\mathcal{U}(0.2, 0.4)$
Infiltration Vide Sanitaire COMBLE Etage RDJ	infiltration infiltration infiltration infiltration	infil.vs infil.comble infil.etage Infil.rdj	$\mathcal{U}(1.8, 2.2)$ $\mathcal{U}(1.8, 2.2)$ $\mathcal{U}(0.1035, 0.1265)$ $\mathcal{U}(0.1035, 0.1265)$



# Résumé

Ce travail de thèse s'inscrit dans le cadre du projet ANR FIABILITE qui porte sur la fiabilité des logiciels de simulation thermique dynamique et plus particulièrement sur les sources potentielles de biais et d'incertitude dans le domaine de la modélisation thermique et énergétique des bâtiments basse consommation. Les sollicitations telles que les occupants, la météo ou encore les scénarios de consommation des usages font partie des entrées les plus incertaines et potentiellement les plus influentes sur les performances d'un bâtiment basse consommation. Il est nécessaire pour pouvoir garantir des performances de déterminer les dispersions de sortie associées à la variabilité des entrées temporelles et d'en déterminer les variables responsables pour mieux réduire leur variabilité ou encore concevoir le bâtiment de manière robuste.

Pour répondre à cette problématique, on se base sur les indices de sensibilité de Sobol adaptés aux modèles complexes à grandes dimensions tels que les modèles de bâtiment pour la simulation thermique dynamique. La gestion des entrées fonctionnelles étant un verrou scientifique pour les méthodes d'analyse de sensibilité standard, une méthodologie originale a été développée dans le cadre de cette thèse afin de générer des échantillons compatibles avec l'estimation de la sensibilité. Bien que la méthode soit générique aux entrées fonctionnelles, elle a été validée dans ce travail de thèse pour le cas des données météorologiques et tout particulièrement à partir des fichiers météo moyens (TMY) utilisés en simulation thermique dynamique.

Les deux aspects principaux de ce travail de développement résident dans la caractérisation de la variabilité des données météorologiques et dans la génération des échantillons permettant l'estimation de la sensibilité de chaque variable météorologique sur la dispersion des performances d'un bâtiment. A travers différents cas d'application dérivés du modèle thermique d'une maison basse consommation, la dispersion et les paramètres influents relatifs à la variabilité météorologique sont estimés. Les résultats révèlent un intervalle d'incertitude sur les besoins énergétiques de l'ordre de 20% à 95% de niveau de confiance, dominé par la température extérieure et le rayonnement direct.

**Mots clés :** Analyse de sensibilité et d'incertitude, entrées temporelles, variabilité météorologique, bâtiment basse consommation, simulation thermique dynamique, fiabilité.



# Abstract

This thesis is part of the ANR project FIABILITE dealing with the reliability of dynamic thermal simulation softwares and particularly with the potential sources of bias and uncertainties in the field of thermal and energy modeling of low consumption buildings. The solicitations such as the occupancy schedules, the weather data or the usage scenarios are among the most uncertain and potentially most influential inputs on the performance of a low energy building. To ensure the efficiency of such buildings, we need to determine the outputs dispersion associated with the uncertainty of the temporal inputs as well as to emphasize the variables responsible for the dispersion of the output in order to design the building in a robust manner.

To address this problem, we have used the sensitivity indices of Sobol adapted to complex models with high dimensions, such as building models for dynamic thermal simulations. The management of the functional inputs being a lock for the scientific methods of standard sensitivity analysis, an innovative methodology was developed in the framework of this thesis in order to generate consistent samples with the estimate of the sensitivity. Although the method can incorporate generic functional inputs, it has been validated in this thesis using meteorological data and especially the typical meteorological year (TMY files) used in dynamic thermal simulations.

The two main aspects of this development work lie in the characterization of the variability of meteorological data and the generation of samples to estimate the sensitivity of each weather variable dispersion on the thermal and energy performances of a building. Through various case studies derived from the thermal model of a low-energy house, the dispersion and influential parameters for meteorological variability are estimated. Results show a large range of uncertainties in the energy requirements from about 20 % at a confidence level of 95 %.

**Keywords** : Uncertainty and sensitivity analyses, temporal inputs, meteorological variability, low-consumption building, dynamic thermal simulation, reliability.